Université de Provence, Aix Marseille I
U.F.R. de Mathématiques, Informatique et Mécanique

Numéro : | | | | | | | | | | |

Thèse

pour obtenir le grade de

Docteur de l'Université de Provence.

Discipline : Mathématiques.

*Ecole Doctorale en Mathématiques
et Informatique de Marseille.*

Présentée et soutenue publiquement par

Anthony MICHEL

le 26 octobre 2001.

Convergence de schémas volumes finis pour des problèmes de convection diffusion non linéaires

Directeurs de thèse :

Robert EYMARD et **Raphaèle HERBIN**

Jury :

Marc DURAND	Professeur, Université de Provence	*Président*
Fayssal BENKHALDOUN	Professeur, Université Paris 13	*Examinateur*
Yann BRENIER	Professeur, détaché C.N.R.S.	*Rapporteur*
Michel CROUZEIX	Professeur, Université de Rennes	*Rapporteur*
Robert EYMARD	Professeur, Université de Marne la Vallée	
Raphaèle HERBIN	Professeur, Université de Provence	

à Julien...

Remerciements

Je tiens à remercier en premier lieu les membres du jury devant lesquels j'ai eu l'honneur de présenter ma thèse le vendredi 26 octobre 2001.

Merci à Fayssal Benkhaldoun avec qui j'ai souvent eu l'occasion de discuter des méthodes de volumes finis. C'est toujours un plaisir de le retrouver dans les colloques d'analyse numérique.

Merci à Marc Durand, professeur autodidacte passionné. J'apprécie à sa juste mesure l'intérêt constant qu'il a eu pour mon travail même si son domaine de prédilection n'est pas précisément le mien.

Merci enfin à Michel Crouzeix et Yann Brenier qui ont accepté d'être rapporteurs sur ma thèse. Ils ont tous deux su parfaitement comprendre les enjeux de mes travaux et je voudrais qu'ils sachent que j'ai apprécié leurs remarques constructives.

Je remercie sincèrement mes directeurs, Robert Eymard et Raphaèle Herbin. Merci de m'avoir proposé ce sujet de thèse passionnant aussi bien du point de vue l'étude des équations aux dérivées partielles que de l'analyse numérique. Merci de m'avoir fait découvrir le métier d'enseignant chercheur dans des conditions aussi favorables. J'espère qu'un jour j'arriverai à suivre votre exemple en matière de pédagogie, de curiosité, de ténacité, de rigueur et de travail d'équipe.

J'ai rejoint l'équipe de Thierry Gallouët à Marseille en quatrième année de l'école Normale Supérieure de Cachan. J'ai reçu un excellent accueil de tous les membres, permanents ou non, et je suis reconnaissant envers Jean Michel Ghidaglia, ancien directeur du département de mathématiques de Cachan, pour ses conseils avisés.

A la manière des phénomènes que j'ai étudiés dans ma thèse, l'évolution des travaux de recherche est vraiment non linéaire en pratique, et même parfois discontinue. Pour ma part, j'ai toujours eu la chance de pouvoir discuter de mes problèmes avec d'autres chercheurs, au sein de l'équipe ou lors des séminaires auxquels j'étais invité à participer et cela m'a permis de réduire au maximum les discontinuités. Je ne pourrai pas tous les citer ici, mais s'ils se reconnaissent, qu'ils sachent qu'ils ont tous contribué à leur manière à la construction de ma thèse.

Deux d'entre eux ont accepté de m'accompagner dans mes recherches. Travailler à leur côté a été un grand plaisir pour moi. Il s'agit de Mario Ohlberger, de l'Université de Freiburg en Allemagne, et de Julien Vovelle qui termine actuellement sa thèse à Marseille. Je tiens à les remercier en particulier et j'espère que nous continuerons encore longtemps à collaborer.

Mes remerciements les plus affectueux vont à mes parents qui m'ont toujours soutenu dans mes choix et à ma femme Céline qui a toujours été attentive à mon travail et qui n'a cessé de m'encourager au cours de ces trois dernières années.

Introduction générale

Ce mémoire est centré autour de l'analyse numérique des schémas volumes finis pour un modèle simplifié d'écoulement de deux fluides incompressibles en milieu poreux. De tels modèles proviennent par exemple des écoulements eau huile dans les gisements d'hydrocarbures en ingénierie pétrolière ou des écoulements eau-air dans les nappes phréatiques en hydrologie. Ces phénomènes sont souvent qualifiés de phénomènes de convection diffusion à convection dominante ("convection dominated problems" en anglais) car si les deux phénomènes sont bien présents, l'ordre de grandeur des données physiques tend à rendre le terme de diffusion très petit devant le terme de convection. Cependant chacun des termes a son importance et doit être pris en compte aussi bien au niveau de la physique que dans l'étude mathématique.

La première partie de cette thèse est consacrée à l'approximation numérique d'équations paraboliques hyperboliques faiblement ou fortement dégénérées. Nous nous sommes intéressés à ces équations pour plusieurs raisons. En premier lieu, dans les écoulements que nous étudions, la modélisation débouche sur un système couplé entre la pression et la saturation et l'équation en saturation est effectivement une équation de type parabolique hyperbolique avec un terme de diffusion pouvant dégénérer. En second lieu, il était intéressant d'adapter les résultats récents de José Carrillo [Car99], sur l'unicité des solutions entropiques pour ce type d'équations, à l'étude des schémas volumes finis. Les trois premiers chapitres sont consacrés à l'étude de la convergence de schémas volumes finis. Le dernier chapitre est consacré à l'analyse des résultats numériques obtenus.

La seconde partie est consacrée à l'analyse numérique d'un modèle simplifié d'écoulement diphasique en milieu poreux. Nous utilisons essentiellement les outils et méthodes développés dans la première partie ou antérieurement (cf [EGH00b]) pour étudier la convergence de deux schémas volumes finis différents pour le phénomène physique complet. Dans les deux cas, l'étude est basée sur des estimations a priori dont l'adaptation au cas discret pose de nombreuses difficultés. Le premier schéma dit "des mathématiciens" est basé sur la réécriture du système (4.2)(4.3) sous la forme d'une équation parabolique hyperbolique sur la saturation et d'une équation elliptique sur la pression, ces deux équations étant couplées par le coefficient de diffusion. Le second schéma dit schéma "des pétroliers" est une méthode numérique utilisée en pratique dans l'industrie pétrolière. L'idée est de faire un décentrage des équations phase par phase. Les deux schémas sont analysés séparément et ils sont ensuite comparés numériquement.

A l'exception des résultats numériques et des preuves données en annexe, les différents paragraphes sont des articles qui ont été soit soumis à des revues mathématiques, soit publiés dans des actes de congrès avec comité de lecture. Ces articles rédigés en anglais peuvent être lus indépendamment les uns des autres, même si l'ordre dans lequel ils sont présentés suit la chronologie de mes travaux de recherche. Ils sont précédés d'un résumé en français destiné à replacer chaque travail dans son contexte.

Table des Matières

Remerciements **3**

Introduction générale **5**

I Convergence de schémas volumes finis pour les équations paraboliques hyperboliques dégénérées **11**

Introduction partie I **13**

1 Equation parabolique hyperbolique faiblement dégénérée avec condition de Neumann homogène **15**

1.1 The problem . . . 17
1.2 The scheme . . . 17
1.3 Discrete estimates on $\varphi(u)$ and $f(u)$ . . . 18
1.4 Translates estimates . . . 20
1.5 Compactness and Convergence . . . 21
1.6 Uniqueness . . . 23

2 Equation parabolique hyperbolique dégénérée avec condition de Dirichlet non homogène dans le cas $\mathbf{q} \cdot \mathbf{n} = 0$ **25**

2.1 The nonlinear parabolic degenerate problem. . . . 27
2.2 Finite volume approximation and main convergence result . . . 29
2.3 Existence, uniqueness and discrete properties . . . 32
2.4 Compactness of a family of approximate solutions . . . 39
2.5 Convergence towards an entropy process solution . . . 42
2.6 Uniqueness of the entropy process solution. . . . 48
2.7 Conclusion . . . 56
2.8 A numerical example . . . 57

3 Equation parabolique hyperbolique dégénérée avec condition de Dirichlet non homogène dans le cas général **59**

3.1 Introduction . . . 61
3.2 Entropy weak solution . . . 63
3.3 Entropy process solution . . . 65
3.4 Uniqueness of the entropy process solution . . . 67
3.4.1 Proof of Theorem 3.4.1, definitions and notations . . . 67

3.4.2 First step in the proof of Theorem 3.4.1: study of the behavior of an entropy process solution near the boundary . . . 69
3.4.3 Second step in the proof of Theorem 3.4.1: inner comparison . . . 72
3.4.4 Third step in the proof of Theorem 3.4.1: the doubling variable method . . . 73
3.5 The finite volume scheme . . . 81
3.5.1 Assumptions and notations . . . 81
3.5.2 The scheme . . . 83
3.6 Monotonicity of the scheme and direct consequences . . . 83
3.7 A priori estimates . . . 86
3.8 Continuous entropy inequalities . . . 90
3.9 Convergence of the scheme . . . 96
3.9.1 Non-linear weak star compactness . . . 96
3.9.2 Compactness in $L^2(Q)$ . . . 96

4 Résultats numériques 99
4.1 Test parabolique dégénéré 1D . . . 99
4.2 Test parabolique dégénéré 2D . . . 101

II Analyse numérique d'un modèle simplifié d'écoulement diphasique incompressible en milieu poreux 105

Introduction partie II 107

Justification du modèle 109

5 Convergence du schéma volumes finis des "mathématiciens" 111
5.1 Introduction . . . 113
5.2 The finite volume scheme . . . 115
5.2.1 Definitions and notations . . . 115
5.2.2 The scheme . . . 116
5.2.3 *A priori* estimates . . . 117
5.2.4 Existence of a solution to the scheme . . . 120
5.3 Convergence results . . . 121
5.3.1 Compactness of $u_{\mathcal{T},\delta t}$ . . . 121
5.3.2 Compactness of $\theta_{\mathcal{T},\delta t}$ . . . 124
5.3.3 Convergence theorem . . . 124
5.4 Numerical results . . . 127

6 Convergence d'un schéma volumes finis pour un écoulement diphasique en milieu poreux avec un décentrement phase par phase 129
6.1 Introduction . . . 131
6.2 The finite volume scheme . . . 134
6.2.1 Finite volume definitions and notations . . . 134
6.2.2 The coupled finite volume scheme . . . 135
6.3 Discrete a priori estimates . . . 137
6.3.1 The maximum principle . . . 137
6.3.2 Estimates on the pressure . . . 138
6.3.3 Existence of a discrete solution . . . 143

6.3.4 Estimates on $g(u)$ . . . 144
6.4 Compactness properties . . . 150
6.5 Study of the limit . . . 154
6.6 Preuve de l'inégalité (6.1.10) . . . 157

7 Résultats numériques 159
7.1 Comparaison des deux schémas en 1D . . . 159
7.1.1 Test comparatif (système découplé) . . . 159
7.1.2 Test en absence de pression capillaire . . . 161
7.1.3 Test avec couplage des équations . . . 161
7.2 Calcul en 2D par le schéma "des pétroliers" . . . 163

Conclusion 165
Résultats . . . 165
Perspectives . . . 166

Annexes 167

A Unicité des solutions faibles pour les équations paraboliques faiblement dégénérées par dualité 169
A.1 Définition de la notion de solution . . . 169
A.2 Unicité par la méthode duale . . . 170
A.2.1 Principe de la démonstration d'unicité . . . 170
A.2.2 Théorème d'existence pour un problème proche . . . 170
A.2.3 Principe du maximum et comparaison . . . 171
A.2.4 Une estimation sur $\Delta\psi$ et $\nabla\psi$ . . . 171
A.2.5 Résultat d'unicité et démonstration . . . 173
A.3 Démonstration du théorème A.2.2 . . . 174

Bibliographie 183

Partie I

Convergence de schémas volumes finis pour les équations paraboliques hyperboliques dégénérées

Introduction

Cette partie est consacrée à l'étude du problème d'évolution suivant :

$$u_t - \Delta\varphi(u) + \mathrm{div}(\mathbf{q}f(u)) = 0, \tag{1}$$

dans un domaine borné et avec une condition initiale u_0 essentiellement bornée. L'équation (1) est une équation de type parabolique hyperbolique non linéaire. Lorsque φ vérifie l'hypothèse $\varphi' > \alpha > 0$, on dit qu'il s'agit d'une équation parabolique non linéaire non dégénérée. Le comportement de la solution est alors quasiment le même que celui de la solution de l'équation de la chaleur, en particulier, on peut montrer dans ce cas l'existence et l'unicité d'une solution faible au sens des distributions (cas particulier de la preuve donnée dans l'annexe A). Lorsque φ' peut s'annuler à l'intérieur de l'intervalle défini par les bornes de u_0 (et de la condition au bord pour des conditions de type Dirichlet), il n'en va pas de même. En effet, si le terme $\Delta\varphi(u)$ s'annule, le problème étudié dégénère en une équation hyperbolique non linéaire et il est bien connu que ce problème est mal posé au sens des distributions. Pour conserver l'unicité on doit alors ajouter des conditions supplémentaires, aussi appelées conditions d'entropie, introduites par Kruzkov [Kru70] dans les années 60.

Nous avons traité trois cas dégénérés ou fortement dégénérés, chacun correspondant à des conditions au bord différentes.

Le premier article intitulé “Convergence of finite volumes methods for convection diffusion problems” traite le cas faiblement dégénéré, c'est à dire le cas où l'ensemble $\{x \in \mathbb{R}, \varphi'(x) = 0\}$ est de mesure nulle, avec des conditions au bord de type Neumann homogène. Dans ce cas, la formulation faible suffit encore pour assurer l'unicité. Cet article a fait l'objet d'une présentation orale au congrès “Finite volume methods for complex applications II” qui s'est tenu à Duisburg en 1999 et a été publié dans les actes du congrès [Mic99]. Cet article s'inspire en grande partie du travail de Younes Naït Slimane [EGHNS98][NS97], étudiant en thèse de Robert Eymard décédé en 1995.

Le second article intitulé “Convergence of finite volumes methods for parabolic degenerate problems” traite le cas fortement dégénéré où l'ensemble $\{x \in \mathbb{R}, \varphi' = 0\}$ est de mesure quelconque, avec des conditions au bord de type Dirichlet non homogènes. La partie numérique est un prolongement de l'article précédent et en même temps un trait d'union entre les méthodes de démonstration utilisées pour les équations elliptiques (cf [EGH99]) et celles utilisées pour les équations hyperboliques (cf [CH99]). La démonstration d'unicité s'inspire des travaux de José Carrillo [Car99] sur les solutions entropiques pour les équations paraboliques. Le traitement des conditions au bord de Dirichlet non homogènes est rendu possible par l'hypothèse que nous faisons sur $\mathbf{q}$, à savoir $\mathbf{q} \cdot \mathbf{n} = 0$. En cela il ne s'agit ni d'un cas particulier, ni d'une extension des travaux précédents qui traitent le cas homogène, mais bien d'une adaptation. D'autre part, le cas envisagé peut être interprété dans le cadre de la modélisation des écoulements diphasiques. La méthode utilisée pour l'existence passe par la notion de solution mesure (appelée ici processus entropique) introduite par Di Perna [DiP85][EGGH98] pour l'étude des équations hyperboliques. Cet outil permet d'obtenir la convergence faible de $g(u_n)$, pour une suite u_n de fonctions bornée dans L^∞ et toute fonction g continue bornée, au prix de l'ajout d'un paramètre que la preuve d'unicité permet ensuite d'éliminer. Cet article a été accepté pour publication dans la revue Numerische Mathematik.

Le troisième article est un travail commun avec Julien Vovelle, qui au moment de ma soutenance finira sa deuxième année de thèse au LATP. Nous donnons dans cet article une formulation intégrale de la notion de solution entropique pour des conditions aux limites de type Dirichlet non homogènes dans le cas général $\mathbf{q} \cdot \mathbf{n} \neq 0$. Le cas traité est d'ailleurs plus général car notre étude concerne les équations du type $u_t - \Delta\varphi(u) + \mathrm{div}(F(x,t,u)) = 0$, avec des hypothèses raisonnables sur F. Pour la partie spécifique à la convergence du schéma numérique nous avons repris les arguments du précédent article mais les conditions aux bord sont traitées avec beaucoup plus d'attention. La preuve de l'unicité s'inspire quand à elle des travaux de Félix Otto [Ott96a][MNRR96] et de Julien Vovelle [Vov00].

L'avancée réalisée au cours de ces travaux est enthousiasmante pour plusieurs raisons : tout d'abord, les techniques utilisées pour les équations elliptiques et hyperboliques ont été mêlées avec succès dans le cas des équations paraboliques comme pour l'étude de systèmes couplés modélisant les écoulements diphasiques en milieu poreux (Partie II). Ensuite, la convergence des schémas volumes finis nous donne l'existence d'une solution à des problèmes encore difficiles à appréhender en fournissant un méthode efficace pour l'approcher. Les résultats numériques (Chapitre 4) permettent également de mieux comprendre les conditions aux limites et le comportement des solutions. Enfin, le fait d'aborder la question sous l'angle des volumes finis nous a obligés à formuler le problème sous forme intégrale avec des fonctions simples, comme les demi entropies de Kruzkov (Formules (3.3)). Nous avons ainsi obtenu une notion de solution entropique qui étend la formulation donnée par José Carrillo [Car99] pour des conditions de Dirichlet non homogènes.

Chapitre 1

Equation parabolique hyperbolique faiblement dégénérée avec condition de Neumann homogène

Résumé

Dans cet article, on construit un schéma volumes finis explicite pour une équation de convection diffusion non linéaire. Ce schéma vérifie un principe du maximum sous une condition habituelle de type C.F.L. . On montre également que la solution approchée vérifie des estimations a priori pour une norme discrète correspondant à la norme $L^2(0,T,H^1(\Omega))$. Ces estimations suffisent pour appliquer le théorème de Fréchet-Kolmogorov. On obtient alors une convergence faible du schéma, à une sous-suite extraite près. Dans la dernière partie, on donne les éléments d'une preuve d'unicité par une méthode duale dans un cas un peu moins général (la preuve complète est donnée en annexe A).

Ce travail s'inspire de la thèse de Younès Nait Slimane [EGHNS98][NS97] qui a démontré une convergence faible en utilisant la notion de solution processus entropique. En reprenant la même méthodologie, on obtient une preuve directe de la convergence forte adaptée au cas faiblement dégénéré envisagé ici. Ce travail a permis de montrer que le traitement du cas faiblement dégénéré pouvait se faire simplement avec les méthodes déjà connues, ce qui permet notamment de mieux comprendre les différences qui existent avec le cas fortement dégénéré. Il a fait l'objet d'une présentation orale au congrès "Finite Volumes for Complex applications II" [Mic99].

Convergence of a finite volume scheme for a nonlinear convection-diffusion problem

Abstract

We construct a time explicit scheme for a nonlinear convection diffusion problem which is L^∞ stable under a C.F.L. condition. We also obtain discrete estimates that allow us to apply Kolmogorov's theorem. Then up to a subsequence we get the convergence of the scheme to a weak solution of our problem. In the last section, which is a summary of Annex A, we give some elements of a simple proof of uniqueness in a special case by a dual method.

1.1 The problem

We consider the following nonlinear parabolic degenerate problem in a bounded polygonal domain Ω of $\mathbb{R}^q$:

$$\begin{cases} u_t - \Delta\varphi(u) + \mathrm{div}(\mathbf{q}f(u)) = 0 \text{ in } \Omega \times (0,T) \\ u(\cdot,0) = u_0(\cdot) \text{ in } \Omega \\ \nabla\varphi(u)\cdot\mathbf{n} = 0 \text{ in } \partial\Omega \times (0,T) \end{cases} \tag{1.1}$$

This problem arises in the study of a two-phase flow in a porous medium. The unknown u is the saturation of the first phase so it takes its values in the interval $[0,1]$. The global flux vector $\mathbf{q}$ is a given function in $\mathcal{C}^1(\bar\Omega \times [0,T])$. In the same way, f and φ are given in $\mathcal{C}^1([0,1])$. The only assumptions that we make are

- (H1) $\mathrm{div}(\mathbf{q}) = 0$ in $\Omega \times (0,T)$
- (H2) $\mathbf{q}\cdot\mathbf{n} = 0$ in $\partial\Omega \times (0,T)$
- (H3) φ is strictly increasing but φ' may vanish (in which case the equation degenerates) and f is nondecreasing.

1.2 The scheme

Let $\mathcal{T}$ be a cell-centered unstructured admissible mesh (see Definition 2.2.1, Chapter 2 and δt a time step. We will always suppose for the sake of simplicity that there exist $N \in \mathbb{N}$ such that $(N+1)\delta t = T$. We construct an approximate piecewise constant solution a.e. on $\Omega \times [0,T)$ by

$$u_{\mathcal{T},\delta t}(x,t) = u_K^n \text{ if } (x,t) \in K \times [n\delta t, (n+1)\delta t),$$

where $\{u_K^n\}$ is defined by an "upwind" finite volume scheme which we now describe.

For a control volume $K \in \mathcal{T}$ we take

$$u_K^0 = \frac{1}{m(K)}\int_K u_0(x)dx.$$

For $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$ we define u_K^{n+1} from $\{u_K^n\}_{K\in\mathcal{T}}$ by

$$m(K)\frac{u_K^{n+1} - u_K^n}{\delta t} - \sum_{L\in\mathcal{N}(K)} T_{K,L}(\varphi(u_L^n) - \varphi(u_K^n)) + \sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^n f(u_{K,L}^{n\,\oplus}) = 0, \tag{1.2}$$

where $\mathcal{N}(K)$ is the set of neighbor control volumes of K, $\mathbf{q}_{K,L}^n = \int_\sigma \mathbf{q}(\cdot, n\delta t)\cdot \mathbf{n}_{K,L}$ if $\sigma = \bar{K}\cap\bar{L}$, $T_{K,L} = \frac{m(\bar{K}\cap\bar{L})}{d(x_K,x_L)}$ and $u_{K,L}^{n\,\oplus}$ is an upwind choice of u on the interface between K and L (it equals u_K^n if $\mathbf{q}_{K,L}^n \geq 0$ and u_L^n otherwise). As a consequence of Hypotheses (H1)(H2), we have

Proposition 1.2.1 (Discrete divergence free property)

$$\sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^n = 0.$$

If we denote by $x^- = \max(0, -x)$, Equation (1.2) is therefore equivalent to the following non-conservative form

$$m(K)\frac{u_K^{n+1} - u_K^n}{\delta t} - \sum_{L\in\mathcal{N}(K)} T_{K,L}(\varphi(u_L^n) - \varphi(u_K^n)) + \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- (f(u_K^n) - f(u_L^n)) = 0. \tag{1.3}$$

By using monotony of the scheme we then easily prove the following lemma

Lemma 1.2.1 **(L^∞ stability)** *Let $\Phi = \max_{x\in[0,1]} |\varphi'(x)|$ and $F = \max_{x\in[0,1]} |f'(x)|$. Assume that*

$$\frac{\delta t}{m(K)} \sum_{L\in\mathcal{N}(K)} T_{K,L}\Phi + {\mathbf{q}_{K,L}^n}^- F \leq 1 \text{ for all } K \in \mathcal{T}.$$

Then the scheme is L^∞ stable, i.e. if $u_0 \in [A, B]$ almost everywhere then $u_{\mathcal{T},\delta t} \in [A, B]$ a.e. .

1.3 Discrete estimates on $\varphi(u)$ and $f(u)$

Lemma 1.2.1 gives an L^∞ estimate on $u_{\mathcal{T},\delta t}$. But it is not enough to deal with the nonlinear terms. We now prove the following inequalities which are crucial for the proof of convergence.

Proposition 1.3.1 *Let $\Phi = \max_{x\in[0,1]} |\varphi'(x)|$ and $F = \max_{x\in[0,1]} |f'(x)|$. On condition that*

$$\frac{\delta t}{m(K)} \sum_{L\in\mathcal{N}(K)} T_{K,L}\Phi + {\mathbf{q}_{K,L}^n}^- F \leq 1 - \varepsilon < 1 \text{ for all } K \in \mathcal{T}, \tag{1.4}$$

then there exist $C(\xi, u_0) > 0$ such that

$$\sum_{n=0}^{N} \delta t \, \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K,L}(\varphi(u_K^n) - \varphi(u_L^n))^2 \leq C\Phi, \tag{1.5}$$

and

$$\sum_{n=0}^{N} \delta t \, \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- (f(u_K^n) - f(u_L^n))^2 \leq CF. \tag{1.6}$$

Proof. We use again the non conservative form (1.3). We multiply the equation by $\delta t u_K^n$ and sum over $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$. We get $E_1 + E_2 + E_3 = 0$, where

$$\begin{aligned}
E_1 &= \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} m(K)(u_K^{n+1} - u_K^n)u_K^n \\
&= \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} m(K) \left(\frac{1}{2}(u_K^{n+1})^2 - \frac{1}{2}(u_K^n)^2 - \frac{1}{2}(u_K^{n+1} - u_K^n)^2 \right) \\
&= \frac{1}{2} \sum_{K\in\mathcal{T}} m(K)(u_K^{N+1})^2 - \frac{1}{2} \sum_{K\in\mathcal{T}} m(K)(u_K^0)^2 - \frac{1}{2} \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} m(K)(u_K^{n+1} - u_K^n), \\
E_2 &= \sum_{n=0}^{N} \delta t \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K,L}(\varphi(u_K^n) - \varphi(u_L^n))u_K^n \\
&= \sum_{n=0}^{N} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K,L}(\varphi(u_K^n) - \varphi(u_L^n))(u_K^n - u_L^n) \\
&\geq \sum_{n=0}^{N} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K,L}(\varphi(u_K^n) - \varphi(u_L^n))^2 \frac{1}{\Phi}
\end{aligned}$$

and $\quad E_3 = \sum_{n=0}^{N} \delta t \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- \, (f(u_K^n) - f(u_L^n))u_K^n.$

Let give now a useful result (see for instance [EGH00b]).

Lemma 1.3.1 (Technical lemma) *Let $g(x) = \int_\alpha^x t f(t)dt$ where $f \in \mathcal{C}^1([\alpha, \beta])$ is* monotone *then if we denote by $F = \max_{x\in[\alpha,\beta]} |f'(x)|$, for all a and $b \in [\alpha, \beta]$,*

$$(f(b) - f(a))b \geq g(b) - g(a) + \frac{1}{2F}(f(b) - f(a))^2.$$

By lemma 1.3.1,

$$\begin{aligned}
E_3 \;\geq\; & \sum_{n=0}^{N} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- \, (f(u_K^n) - f(u_L^n))^2 \frac{1}{F} \\
& + \sum_{n=0}^{N} \delta t \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- \, (g(u_K^n) - g(u_L^n)).
\end{aligned}$$

But $\mathbf{q}_{K,L}^n = -\mathbf{q}_{L,K}^n$. So by using Proposition 1.2.1, we get

$$\sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- \, (g(u_K^n) - g(u_L^n)) = 0$$

and finally

$$E_3 \;\geq\; \sum_{n=0}^{N} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} {\mathbf{q}_{K,L}^n}^- \, (f(u_K^n) - f(u_L^n))^2 \frac{1}{F}.$$

It remains to estimate $R_1 = \sum_{n=0}^{N} \frac{1}{2} \sum_{K\in\mathcal{T}} m(K)(u_K^{n+1} - u_K^n)^2$. By 1.3,

$$
\begin{aligned}
(u_K^{n+1} - u_K^n)^2 \leq & \left(\frac{\delta t}{m(K)}\right)^2 \left(\sum_{L\in\mathcal{N}(K)} T_{K,L}|\varphi(u_K^n) - \varphi(u_L^n)| + {\mathbf{q}_{K,L}^n}^- |f(u_K^n) - f(u_L^n)| \right)^2 \\
\leq & \left(\frac{\delta t}{m(K)}\right)^2 \left(\sum_{L\in\mathcal{N}(K)} T_{K,L}\Phi + {\mathbf{q}_{K,L}^n}^- F \right) \Bigg(\sum_{L\in\mathcal{N}(K)} \frac{1}{\Phi} T_{K,L}(\varphi(u_K^n) - \varphi(u_L^n))^2 \\
& + \sum_{L\in\mathcal{N}(K)} \frac{1}{F} {\mathbf{q}_{K,L}^n}^- (f(u_K^n) - f(u_L^n))^2 \Bigg).
\end{aligned}
$$

Therefore by using Condition (1.4), we get

$$
R_1 \leq (1-\varepsilon) \sum_{n=0}^{N} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \frac{1}{\Phi} T_{K,L}(\varphi(u_K^n) - \varphi(u_L^n))^2 + \frac{1}{F} {\mathbf{q}_{K,L}^n}^- (f(u_K^n) - f(u_L^n))^2.
$$

Moreover, $\sum_{K\in\mathcal{T}} m(K)(u_K^0)^2 \leq \|u_0\|_{L^2(\Omega)}^2$. Then by collecting the previous inequalities, we complete the proof with $C(\xi, u_0) = \frac{1}{2}\frac{\|u_0\|_{L^2(\Omega)}^2}{1-\varepsilon}$.
□

1.4 Translates estimates

As a consequence of Estimate (1.5), by classical techniques that can be seen for instance in [EGH00b], we get the two following inequalities.

Corollary 1.4.1 (Space-translate estimate)
Under the assumptions of Proposition 1.3.1,

$$
\int_0^T \int_{\Omega_\xi} \left(\varphi(u_{\mathcal{T},\delta t}(x+\xi, t)) - \varphi(u_{\mathcal{T},\delta t}(x,t))\right)^2 dxdt \leq C\Phi|\xi|(|\xi| + 2h),
$$

where $h = \max_{K\in\mathcal{T}} \operatorname{diam}(\mathrm{K})$ *and* $\Omega_\xi = \Omega \cap \Omega - \xi$.

Corollary 1.4.2 (Time-translate estimate)
Under the assumptions of Proposition 1.3.1, there exist $C'(\varepsilon, \varphi, f, \mathbf{q}, u_0, \Omega, T) > 0$ *such that*

$$
\int_0^{T-s} \int_\Omega (\varphi(u_{\mathcal{T},\delta t}(x, t+s)) - \varphi(u_{\mathcal{T},\delta t}(x,t)))^2 dxdt \leq C's.
$$

1.5 Compactness and Convergence

We handle now the first part of the conclusion under the convergence of the scheme. For that we need to use the classical following compactness theorem.

Theorem 1.5.1 (Fréchet-Kolmogorov)
Let $\mathcal{F}$ a bounded subset of $L^2(\mathbb{R}^d)$ such that

$$\lim_{|\xi|\to 0} \sup_{f\in\mathcal{F}} \|f(.+\xi) - f(.)\|_{L^2(\mathbb{R}^d)} = 0.$$

Then for every $\Omega \subset\subset \mathbb{R}^d$, $\mathcal{F}$ is relatively compact in $L^2(\Omega)$.

We are now able to state the convergence theorem

Theorem 1.5.2 (Convergence theorem, part 1) *Let $(\mathcal{T}_m, \delta t_m)$ be a sequence of meshes and time steps that satisfies assumptions of Proposition 1.3.1. Assume moreover that h_m tends to zero when m tends to ∞ (which implies that δt_m tends to zero) .*

Then there exists $u \in L^\infty(\Omega \times (0,T))$ such that $\varphi(u) \in L^2(0,T,H^1(\Omega))$ and up to a subsequence,

$$\lim_{m\to\infty} u_{\mathcal{T}_m,\delta t_m} = u \text{ for the } L^\infty(\Omega\times(0,T)) \text{ weak-}\star \text{ topology and in } L^p(\Omega\times(0,T)),\ \forall p<\infty.$$

Elements of proof. Let us extend $u_{\mathcal{T},\delta t}$ on $\mathbb{R}^{q+1}$ by zero out of $\Omega \times (0,T)$. From Corollary 1.4.1 and Corollary 1.4.2, we directly deduce that for every $(\xi, s) \in \mathbb{R}^{q+1}$,

$$\begin{aligned} \|\varphi(u_{\mathcal{T},\delta t}(\cdot+\xi,\cdot+t)) - \varphi(u_{\mathcal{T},\delta t}(\cdot,\cdot))\|^2_{L^2(\mathbb{R}^{m+1})} &\le 2C|\xi|(|\xi|+2h) + 2C's \\ &+ (2T|\xi|m(\partial\Omega) + 2m(\Omega)s){M_\varphi}^2, \end{aligned}$$

where $M_\varphi = \max\limits_{[0,1]} |\varphi(x)|$.
This inequality allows us to apply Theorem 1.5.1 and we obtain regularity on the limit by looking at the rates of increase which converge to the derivatives in the sense of distributions (see [EGH00b][EGH99])

Theorem 1.5.3 (Convergence theorem, part 2) *We suppose that the assumptions of Theorem 1.5.2 are satisfied, and we assume also that there exists $\theta > 0$ such that for all mesh $\mathcal{T}_m$, the following regularity property is satisfied:*

$$\sum_{L\in\mathcal{N}(K)} m(\bar{K}\cap\bar{L}) \le \theta\frac{m(K)}{h}\ \forall K \in \mathcal{T}. \tag{1.7}$$

Then the function u given in Theorem 1.5.2 is solution of Problem (1.1) in the following sense:
$\forall \psi \in \mathcal{C}_{test} = \{\eta \in \mathcal{C}^{2,1}(\bar{\Omega}\times[0,T])$ such that $\nabla\eta\cdot\mathbf{n} = 0$ and $\eta(\cdot,T) = 0\}$,

$$\int_0^T\int_\Omega u\psi_t + \varphi(u)\Delta\psi + \mathbf{q}f(u)\cdot\nabla\psi \ + \int_\Omega u_0(\cdot)\psi(\cdot,0) = 0. \tag{1.8}$$

Proof. The convergence of u_n to u is strong which implies that $f(u_n)$ and $\varphi(u_n)$ converge to $f(u)$ and $\varphi(u)$. So it suffices to show that

$$\lim_{m\to\infty}\int_0^T\int_\Omega u_m\psi_t+\varphi(u_m)\Delta\psi+\mathbf{q}f(u_m)\cdot\nabla\psi \quad +\int_\Omega u_0(\cdot)\psi(\cdot,0)=0.$$

Let m be fixed and $\mathcal{T}=\mathcal{T}_m$, $\delta t=\delta t_m$. Let us then multiply Equation (1.3) by $\delta t\Psi_K^n$, where $\Psi_K^n=\psi(x_K,n\delta t)$ and sum over $K\in\mathcal{T}$ and $n\in[\![0,N]\!]$. We get $T_1+T_2+T_3=0$, where

$$\begin{aligned}
T_1 &= \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)(u_K^{n+1}-u_K^n)\Psi_K^n\\
&= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}} m(K)u_K^n(\Psi_K^n-\Psi_K^{n-1})+\sum_{K\in\mathcal{T}}u_K^{N+1}\Psi_K^N-\sum_{K\in\mathcal{T}}u_K^0\Psi_K^0,\\
T_2 &= \sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}T_{K,L}(\varphi(u_K^n)-\varphi(u_L^n))\Psi_K^n\\
&= \sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}T_{K,L}(\varphi(u_K^n)-\varphi(u_L^n))(\Psi_K^n-\Psi_L^n)
\end{aligned}$$

$$\text{and}\qquad\begin{aligned}
T_3 &= \sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}\mathbf{q}_{K,L}^{n}{}^{-}\,(f(u_K^n)-f(u_L^n))\Psi_K^n\\
&= \sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}(f(u_K^n)-f(u_L^n))\mathbf{q}_{K,L}^n\Psi_{K,L}^{n\,\ominus},
\end{aligned}$$

where $\Psi_{K,L}^{n\,\ominus}$ is equal to Ψ_K^n if $\mathbf{q}_{K,L}^n\le 0$ and Ψ_L^n otherwise. We compare T_i to S_i given by

$$\begin{aligned}
S_1 &= -\int_0^T\int_\Omega u_m\psi_t-\int_\Omega u_0(\cdot)\psi(\cdot,0)\\
&= -\sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}}u_K^n\frac{1}{\delta t}\int_{n\delta t}^{(n+1)\delta t}\int_K\psi_t(x,t)dxdt-\sum_{K\in\mathcal{T}}\int_K u_0(x)\psi(x,0)dx,\\
S_2 &= \int_0^T\int_\Omega\varphi(u_m)\Delta\psi\\
&= \sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}(\varphi(u_K^n)-\varphi(u_L^n))\frac{1}{\delta t}\int_{n\delta t}^{(n+1)\delta t}\int_{\bar K\cap\bar L}\nabla\psi\cdot\mathbf{n}_{K,L}
\end{aligned}$$

$$\begin{aligned}
\text{and } S_3 &= \int_0^T\int_\Omega\mathbf{q}f(u)\cdot\nabla\psi\\
&= \sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}(f(u_K^n)-f(u_L^n))\frac{1}{\delta t}\int_{n\delta t}^{(n+1)\delta t}\int_{\bar K\cap\bar L}\psi\mathbf{q}\cdot\mathbf{n}_{K,L}.
\end{aligned}$$

Classically, (see for instance [EGHNS98], [EGH00b]) because ψ is a regular function and $\psi(T)=0$ we get $\lim\limits_{m\to\infty}(S_1-T_1)=0$.

By using Estimates (1.6), (1.5) and the regularity condition (1.7) we get also

$$\begin{aligned} |T_3 - S_3| &\leq \sqrt{Tm(\Omega)CFM_{\mathbf{q}}\theta}\ C_2(\psi)\ \sqrt{h} \\ \text{and} \quad |T_2 - S_2| &\leq \sqrt{2Tm(\Omega)C\Phi}\ C_1(\psi)\ h. \end{aligned}$$

Then $\lim_{m\to\infty} |T_2 - S_2| = \lim_{m\to\infty} |T_3 - S_3| = 0$, and the proof is complete. □

1.6 Uniqueness

This section is a summary of the Annex A. Let u_1 et u_2 be two solutions of Problem (1.1) for the weak formulation (1.8). We denote by $u_d = u_1 - u_2$. For all $\psi \in \mathcal{C}_{test}$, we have

$$\int_0^T\int_\Omega u_d(x,t)[\psi_t(x,t) + \mathbf{q}(x,t)F(x,t).\nabla\psi(x,t) + \Phi(x,t)\Delta\psi(x,t)\,]dxdt = 0, \tag{1.9}$$

where $F = \dfrac{f(u_1) - f(u_2)}{u_1 - u_2}$ and $\Phi = \dfrac{\varphi(u_1) - \varphi(u_2)}{u_1 - u_2}$. So it is natural to pay attention to the dual problem:

$$\psi \in C^{2,1}(\bar{\Omega} \times [0,T]) \text{ and } \left\{ \begin{array}{l} \psi_t + \mathbf{q}F.\nabla\psi + \Phi\Delta\psi = \chi \\ \nabla\psi.\mathbf{n} = 0 \\ \psi(T) = 0 \end{array} \right. . \tag{1.10}$$

From [LSU67], we can state the following result

Theorem 1.6.1 (Existence to the regular dual problem)
Let F , $\mathbf{q}$ and Φ be $\mathcal{C}^\infty$ functions under $\bar{\Omega} \times [0,T]$, and assume that there exists $\delta > 0$ such that $\Phi(x,t) \geq \delta$. Then for every $\chi \in \mathcal{C}_c^\infty(\Omega \times (0,T))$ there exists an unique solution to Problem (1.10).

Moreover, we also have the following estimates

Proposition 1.6.1 (Dual problem estimates) *Let ψ be a solution to the regular dual problem with second member χ and M_χ, M_Φ , $M_{\mathbf{q}}$ and M_F some upper bounds for $|\chi|$, Φ, $|\mathbf{q}|$ et $|F|$. Then there exists $C(\chi, M_\Phi, M_{\mathbf{q}}, M_F, \Omega, T) > 0$ such that*

$$\|\psi\|_{L^\infty(\Omega\times(0,T))} \leq M_\chi T \tag{1.11}$$

$$\|\Delta\psi\|_{L^2(\Omega\times(0,T))} \leq \frac{C}{\delta^2}, \tag{1.12}$$

and

$$\|\nabla\psi\|_{L^2(\Omega\times(0,T))} \leq \frac{C}{\delta}. \tag{1.13}$$

Elements of proof. Inequality (1.11) is a direct consequence of the maximum principle for parabolic equations. For (1.12) and (1.13), we multiply the equation by $\Delta\psi$ and integrate over $\Omega \times (0,T)$. Because of (1.11), $\|\nabla\psi\|_{L^2(\Omega\times(0,T))}$ is controlled by $\sqrt{\|\Delta\psi\|_{L^2(\Omega\times(0,T))}}$. We then complete the proof by using time and space integrates by part and Young inequalities.

We are now able to give the main result of this section.

Theorem 1.6.2 (Uniqueness theorem) *Assume that φ^{-1} is a Hölder-continuous function with exponent $\frac{1}{2}$. Then there exists an unique solution to Problem (1.1) for the weak formulation (1.8).*

Proof. By Theorem 1.5.3, there exists at least one solution for the weak formulation (1.8). We now turn to the study of the uniqueness. Let $\chi \in \mathcal{C}_c^\infty(\Omega \times (0,T))$ and $\delta > 0$ ($\delta \leq M_\Phi$).

$\Phi_\delta = \max(\delta, \Phi)$ is again in $L^\infty(\Omega \times (0,T))$ and $\Phi_\delta \geq \delta$ almost everywhere. We don't have regularity hypothesis on Φ_δ and F but we can construct some sequences of regular functions on $\bar{\Omega} \times [0,T]$, (G_n), (F_n) and $(\mathbf{q}_n)$, that converge to Φ_δ, F and $\mathbf{q}$ in $L^p(\Omega \times (0,T))$ for $p < \infty$ and such that

$$\delta \leq G_n \leq M_\Phi,\ |\mathbf{q}_n| \leq M_\mathbf{q} \text{ and } |F_n| \leq M_f.$$

For every n, by Theorem 1.6.1 there exists a solution ψ_n in $C^{2,1}(\bar{\Omega} \times [0,T])$ to the dual problem associated to G_n, $\mathbf{q}_n$, F_n and χ. Because the upper bounds of G_n, $\mathbf{q}_n$, F_n and the lower bound δ of G_n are independent from n, estimates on $\Delta\psi_n$ and $\nabla\psi_n$ are also independent from n, so we get

$$\lim_{\mathbf{n}\to\infty} \int_0^T \int_\Omega u_d(\mathbf{q}_n F_n - \mathbf{q}F).\nabla\psi_n = 0$$

and

$$\limsup_{n\to\infty} \left| \int_0^T \int_\Omega u_d(G_n - \Phi)\Delta\psi_n \right| \leq \left(\int_0^T \int_\Omega u_d^2(\Phi_\delta - \Phi)^2 \right)^{\frac{1}{2}} \frac{C}{\delta^2},$$

But $\Phi_\delta - \Phi \leq \delta\, \mathbb{1}_{\{\Phi<\delta\}}$ because Φ and Φ_δ are equal on $\{\Phi \geq \delta\}$. Then, if we denote by $A_\delta = \{u_d \neq 0\} \cap \{\Phi < \delta\}$, we get

$$\int_0^T \int_\Omega u_d \chi \leq \left(\int_{A_\delta} \frac{u_d^2}{\delta^2} \right)^{\frac{1}{2}}.$$

Because φ^{-1} is a Hölder-continuous function with exponent $\frac{1}{2}$, we get $u_d \leq \delta$ on A_δ. Moreover $m(A_\delta)$ tends to zero, so that

$$\int_0^T \int_\Omega u_d \chi = 0.$$

Since this is true for every regular function χ, the proof is complete. □

Chapitre 2

Equation parabolique hyperbolique dégénérée avec condition de Dirichlet non homogène dans le cas $\mathbf{q} \cdot \mathbf{n} = 0$

Résumé

Nous construisons dans cet article une solution faible entropique approchée pour l'équation non linéaire parabolique dégénérée suivante :

$$u_t + \mathrm{div}(\mathbf{q}\, f(u)) - \Delta\varphi(u) = 0. \tag{2.1}$$

La solution approchée $u_{\mathcal{D}}$ est obtenue par un schéma volumes finis pour une discrétisation admissible $\mathcal{D}$ du domaine qui nous intéresse, à savoir $Q = \Omega \times (0, T)$ où Ω est une domaine borné de $\mathbb{R}^d$. Nous montrons précisément la convergence forte de $u_{\mathcal{D}}$ vers l'unique solution faible entropique de (2.1). Cette preuve est décomposée en deux étapes. Dans un premier temps, grâce à des estimations a priori nous obtenons la convergence faible et à une sous-suite extraite près de $u_{\mathcal{D}}$ vers une solution mesure (appelée solution processus entropique). Dans un deuxième temps nous montrons qu'il existe au plus une solution processus entropique et que cette solution processus entropique se réduit à une fonction de $L^\infty(\Omega \times (0, T))$, c'est à dire que c'est une fonction. Des résultats numériques illustrant le phénomène de dégénérescence parabolique-hyperbolique sont donnés dans la dernière section.

Ce travail commun avec Robert Eymard, Thierry Gallouët et Raphaèle Herbin a permis de montrer comment il était possible d'adapter le travail de José Carrillo [Car99] pour traiter la convergence de méthodes de volumes finis pour des problèmes paraboliques hyperboliques dégénérés sur des domaines bornés. La condition $\mathbf{q} \cdot \mathbf{n} = 0$ que nous avons prise est physiquement acceptable, si on considère que cette équation provient du système couplé étudié dans la Partie II de ce mémoire. D'autre part, la preuve d'unicité avec cette condition limite sur le bord présente déjà de nombreuses difficultés. Nous présentons ici une définition un peu originale de solution entropique et nous montrons que notre formulation est bien adaptée aux schémas volumes finis. Il s'agit du premier article à notre connaissance traitant de l'unicité dans le cas parabolique hyperbolique non linéaire dégénéré avec des conditions au bord non homogènes. Il a été soumis à la fin de l'année 2000 [EGHM00] et j'ai présenté à plusieurs reprises ces travaux lors de séminaires ou de congrès.

Convergence of a finite volume scheme for nonlinear degenerate parabolic equations.

Abstract

One approximates the entropy weak solution u of a nonlinear parabolic degenerate equation $u_t+\mathrm{div}(\mathbf{q}f(u)) -\Delta\varphi(u) = 0$ by a piecewise constant function $u_{\mathcal{D}}$ using a discretization $\mathcal{D}$ in space and time and a finite volume scheme. The convergence of $u_{\mathcal{D}}$ to u is shown as the size of the space and time steps tend to zero. In a first step, estimates on $u_{\mathcal{D}}$ are used to prove the convergence, up to a subsequence, of $u_{\mathcal{D}}$ to a measure valued entropy solution (called here an entropy process solution). A result of uniqueness of the entropy process solution is proved, yielding the strong convergence of $u_{\mathcal{D}}$ to u. Some numerical results on a model equation are shown.

2.1 The nonlinear parabolic degenerate problem.

Let Ω be a bounded open subset of $\mathbb{R}^d$, ($d = 1, 2$ or 3) with boundary $\partial\Omega$ and let $T \in \mathbb{R}_+^*$. One considers the following problem.

$$u_t(x,t) + \mathrm{div}\Big(\mathbf{q}\ f(u)\Big)(x,t) - \Delta\varphi(u)(x,t) = 0,\ \text{for } (x,t) \in \Omega \times (0,T). \tag{2.2}$$

The initial condition is formulated as follows:

$$u(x,0) = u_0(x) \quad \text{for } x \in \Omega. \tag{2.3}$$

The boundary condition is the following non homogenous Dirichlet condition:

$$u(x,t) = \bar{u}(x,t), \quad \text{for } (x,t) \in \partial\Omega \times (0,T). \tag{2.4}$$

This problem arises in different physical contexts. One of them is the problem of two phase flows in a porous medium, such as the air-water flow of hydrological aquifers. In this case, problem (2.2)-(2.4) represents the conservation of the incompressible water phase, described by the water saturation u, submitted to convective flows (first order space terms $\mathbf{q}(x,t)\ f(u)$) and capillary effects ($\Delta\varphi(u)$). The expression $\mathbf{q}(x,t)\ f(u)$ for the convective term in (2.2) appears to be a particular case of the more general expression $F(u,x,t)$, but since it involves the same tools as the general framework, the results of this paper could be extended to some other problems.

One supposes that the following hypothesis, globally referred in the following as hypothesis (H), are fulfilled.

Hypothesis (H)

(H1) Ω *is polygonal (if* $d = 1$, Ω *is an interval, and if* $d = 3$, Ω *is a polyhedron),*

(H2) $u_0 \in L^\infty(\Omega)$ *and* $\bar{u} \in L^\infty(\partial\Omega\times(0,T))$, $\bar{u}$ *being the trace of a function of* $H^1(\Omega\times(0,T))\cap L^\infty(\Omega\times(0,T))$ *(also denoted* $\bar{u}$*); one sets* $U_I = \min(\mathrm{infess}\ u_0, \mathrm{infess}\ \bar{u})$ *and* $U_S = \max(\mathrm{supess}\ u_0, \mathrm{supess}\ \bar{u})$,

(H3) φ *is a nondecreasing Lipschitz-continuous function, with Lipschitz constant* Φ*, and one defines a function* ζ *such that* $\zeta' = \sqrt{\varphi'}$,

(H4) $f \in C^1(\mathbb{R},\mathbb{R})$, $f' \geq 0$*; one sets* $F = \max_{s\in[U_I,U_S]} f'(s)$,

(H5) $\mathbf{q}$ *is the restriction to* $\Omega \times (0,T)$ *of a function of* $C^1(\mathbb{R}^d \times \mathbb{R}, \mathbb{R}^d)$*,*

(H6) $\mathrm{div}(\mathbf{q}(x,t)) = 0$ *for all* $(x,t) \in \mathbb{R}^d \times (0,T)$*, where* $\mathrm{div}(\mathbf{q}(x,t)) = \displaystyle\sum_{i=1}^{d} \frac{\partial q_i}{\partial x_i}(x,t)$*, (*$q_i$ *is the i-eth component of* q*) and*

$$\mathbf{q}(x,t).\mathbf{n}(x) = 0, \quad \textit{for a.e. } (x,t) \in \partial\Omega \times (0,T), \tag{2.5}$$

(for $x \in \partial\Omega$*,* $\mathbf{n}(x)$ *denotes the outward unit normal to* Ω *at point* x*).*

Remark 2.1.1 *The function* f *is assumed to be non decreasing in* $(H3)$ *for the sake of simplicity. In fact, the convergence analysis which we present here would also hold without this monotonicity assumption using for instance a flux splitting scheme for the treatment of the convective term* $\mathbf{q}f(u)$*. However, for general monotonous schemes, the proof is not so clear: the use of the Kruskov entropy pairs which was used for instance in [EGGH98] yields some technical difficulties because of the presence of the second order term.*

Under hypothesis (H), the problem (2.2)-(2.4) does not have, in the general case, strong regular solutions. Because of the presence of a non-linear convection term, the expected solution is an entropy weak solution in the sense of definition 2.1.1 given below.

Definition 2.1.1 (Entropy weak solution) *Under hypothesis (H), a function* u *is said to be an entropy weak solution to Problem (2.2)-(2.4) if it verifies:*

$$u \in L^\infty(\Omega \times (0,T)), \tag{2.6}$$

$$\varphi(u) - \varphi(\bar{u}) \in L^2(0,T; H_0^1(\Omega)), \tag{2.7}$$

and u *satisfies the following Kruzkov entropy inequalities:* $\forall\ \psi \in \mathcal{D}^+(\Omega \times [0,T))$, $\forall \kappa \in \mathbb{R}$,

$$\int_{\Omega\times(0,T)} \left[\begin{array}{l} |u(x,t)-\kappa|\ \psi_t(x,t) + \\ (f(u(x,t)\top\kappa) - f(u(x,t)\bot\kappa))\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t) \\ -\nabla|\varphi(u)(x,t)-\varphi(\kappa)|\cdot\nabla\psi(x,t) \end{array} \right] dxdt + \int_\Omega |u_0(x)-\kappa|\psi(x,0)dx \quad \geq 0, \tag{2.8}$$

where one denotes by $a\top b$ *the maximum value between two real values* a *and* b*, and by* $a\bot b$ *their minimum value and where* $\mathcal{D}^+(\Omega \times [0,T)) = \{\psi \in C_c^\infty(\Omega \times \mathbb{R}, \mathbb{R}_+), \psi(\cdot,T) = 0\}$.

This notion has been introduced by several authors ([Car99], [LBS93]), who proved the existence of such a solution in bounded domains. In [LBS93], the proof of existence uses strong BV estimates in order to derive estimates in time and space for the solution of the regularized problem obtained by adding a small diffusion term. In [Car99], the existence of a weak solution is proved using semi-group theory (see [Ben72]), and the uniqueness of the entropy weak solution is proved using techniques which have been introduced by S.N. Krushkov and extended by J. Carrillo.
In the present study, thanks to condition (2.5), boundary conditions are entirely taken into account by (2.7) and do not appear in the entropy inequality (2.8). For studies of the continuous problem with more general boundary conditions, one can refer to [LBS93], which uses the classical Bardos-Leroux-Nédélec formulation [BlRN79], or [Car99] in the case of a homogeneous Dirichlet boundary condition on $\partial\Omega$ without condition (2.5).

Let us mention some related work in the case of infinite domains ($\Omega = \mathbb{R}^d$): In [BGN00], the authors prove the existence in the case $\Omega = \mathbb{R}^d$, regularizing the problem with the "general kinetic BGK" framework to yield estimates on translates of the approximate solutions. Continuity of the solution with respect to the data for a more general equation was studied by Cockburn and Gripenberg [CG99], and convergence of the discretization with an implicit finite volume scheme was recently studied by Ohlberger [Ohl97].

We shall deal here with the case of a bounded domain. The aim of the present work is then to prove the convergence of approximate solutions obtained using a finite volume method with general unstructured meshes towards the entropy weak solution of (2.2)-(2.4) as the mesh size and time step tend to 0. We state this result in Theorem 2.2.1 in Section 2.2, after presenting the finite volume scheme. Then in Section 2.3, the existence and uniqueness of the solution to the nonlinear set of equations resulting from the finite volume scheme is proven, along with some properties of the discrete solutions. In Section 2.4 we show some compactness properties of the family of approximate solutions. We show in Section 2.5 that there exists some subsequence of sequences of approximate solutions which tends to a so-called "entropy process solution", and in Section 2.6 we prove the uniqueness of this entropy process solution, which allows us to conclude to the convergence of the scheme in Section 2.7. We finally give an example of numerical implementation in Section 2.8.

2.2 Finite volume approximation and main convergence result

Let us first define space and time discretizations of $\Omega \times (0, T)$.

Definition 2.2.1 (Admissible mesh of Ω) *An admissible mesh of Ω is given by a set $\mathcal{T}$ of open bounded polygonal convex subsets of Ω called control volumes, a family $\mathcal{E}$ of subsets of $\bar{\Omega}$ contained in hyper-planes of $\mathbb{R}^d$ with strictly positive measure, and a family of points (the "centers" of control volumes) satisfying the following properties:*

(i) The closure of the union of all control volumes is $\bar{\Omega}$.

(ii) For any $K \in \mathcal{T}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \bar{K} \backslash K = \cup_{\sigma \in \mathcal{E}_K} \bar{\sigma}$. Furthermore, $\mathcal{E} = \cup_{K \in \mathcal{T}} \mathcal{E}_K$.

(iii) For any $(K, L) \in \mathcal{T}^2$ with $K \neq L$, either the "length" (i.e. the $(d-1)$ Lebesgue measure) of $\bar{K} \cap \bar{L}$ is 0 or $\bar{K} \cap \bar{L} = \bar{\sigma}$ for some $\sigma \in \mathcal{E}$. In the latter case, we shall write $\sigma = K|L$ and $\mathcal{E}_{int} = \{\sigma \in \mathcal{E}, \exists (K, L) \in \mathcal{T}^2, \sigma = K|L\}$. For any $K \in \mathcal{T}$, we shall denote by $\mathcal{N}_K$ the set of boundary control volumes of K, i.e. $\mathcal{N}_K = \{L \in \mathcal{T}, K|L \in \mathcal{E}_K\}$.

(iv) The family of points $(x_K)_{K \in \mathcal{T}}$ is such that $x_K \in K$ (for all $K \in \mathcal{T}$) and, if $\sigma = K|L$, it is assumed that the straight line (x_K, x_L) is orthogonal to σ.
For a control volume $K \in \mathcal{T}$, we will denote by $m(K)$ its measure and $\mathcal{E}_{ext,K}$ the subset of the edges of K included in the boundary $\partial\Omega$. If $L \in \mathcal{N}_K$, $m(K|L)$ will denote the measure of the edge between K and L, $\tau_{K|L}$ the "transmissibility" through $K|L$, defined by $\tau_{K|L} = \dfrac{m(K|L)}{d(x_k, x_L)}$. Similarly, if $\sigma \in \mathcal{E}_{ext,K}$, we will denote by $m(\sigma)$ its measure and τ_σ the "transmissibility" through σ, defined by $\tau_\sigma = \dfrac{m(\sigma)}{d(x_K, \sigma)}$. One denotes $\mathcal{E}_{ext} = \cup_{K \in \mathcal{T}} \mathcal{E}_{ext,K}$ and for $\sigma \in \mathcal{E}_{ext}$, one denotes by K_σ the control volume such that $\sigma \in \mathcal{E}_{ext}(K_\sigma)$. The size of the mesh $\mathcal{T}$ is defined by

$$\text{size}(\mathcal{T}) = \max_{K \in \mathcal{T}} \text{diam}(K),$$

and a geometrical factor, linked with the regularity of the mesh, is defined by

$$\text{reg}(\mathcal{T}) = \min_{K \in \mathcal{T}, \sigma \in \mathcal{E}_K} \frac{d(x_K, \sigma)}{\text{diam}(K)}.$$

Remark 2.2.1 *Assumption (iv) in the previous definition is due to the presence of the second order term. Examples of meshes satisfying these assumptions are triangular meshes satisfying the acute angle condition (in fact this condition may be weaken to the Delaunay condition), rectangular meshes or Voronoï meshes, see [EGH99] or [EGH00b] for more details.*

Definition 2.2.2 (Time discretization of $(0,T)$**)** *A time discretization of* $(0,T)$ *is given by an integer value* N *and by an increasing sequence of real values* $(t^n)_{n \in [\![0,N+1]\!]}$ *with* $t^0 = 0$ *and* $t^{N+1} = T$. *The time steps are then defined by* $\delta t^n = t^{n+1} - t^n$, *for* $n \in [\![0,N]\!]$.

Definition 2.2.3 (Space-time discretization of $\Omega \times (0,T)$**)** *A finite volume discretization* $\mathcal{D}$ *of* $\Omega \times (0,T)$ *is the family* $\mathcal{D} = (\mathcal{T}, \mathcal{E}, (x_K)_{K \in \mathcal{T}}, N, (t^n)_{n \in [\![0,N]\!]})$, *where* $\mathcal{T}$, $\mathcal{E}$, $(x_K)_{K \in \mathcal{T}}$ *is an admissible mesh of* Ω *in the sense of definition 2.2.1 and* N, $(t^n)_{n \in [\![0,N+1]\!]}$ *is a time discretization of* $(0,T)$ *in the sense of definition 2.2.2. For a given mesh* $\mathcal{D}$, *one defines:*

$$\text{size}(\mathcal{D}) = \max(\text{size}(\mathcal{T}), (\delta t^n)_{n \in [\![0,N]\!]}), \qquad \textit{and } \text{reg}(\mathcal{D}) = \text{reg}(\mathcal{T}).$$

We may now define the finite volume discretization of (2.2)-(2.4) . Let $\mathcal{D}$ be a finite volume discretization of $\Omega \times (0,T)$ in the sense of Definition 2.2.3. The initial condition is discretized by:

$$U_K^0 = \frac{1}{m(K)} \int_K u_0(x)dx, \quad \forall K \in \mathcal{T}. \tag{2.9}$$

In order to introduce the finite volume scheme, we need to define:

$$\bar{U}_\sigma^{n+1} = \frac{1}{\delta t^n \ m(\sigma)} \int_{t^n}^{t^{n+1}} \int_\sigma \bar{u}(x,t) d\gamma(x) dt, \qquad \forall \sigma \in \mathcal{E}_{ext}, \forall n \in [\![0,N]\!], \tag{2.10}$$

$$q_{K,L}^{n+1} = \frac{1}{\delta t^n} \int_{t^n}^{t^{n+1}} \int_{K|L} \mathbf{q}(x,t) \cdot \mathbf{n}_{K,L} d\gamma(x) dt, \qquad \forall K \in \mathcal{T}, \forall L \in \mathcal{N}_K, \forall n \in [\![0,N]\!], \tag{2.11}$$

where $\mathbf{n}_{K,L}$ be the normal unit vector to $K|L$ oriented from K to L.

An **implicit finite volume scheme** for the discretization of Problem (2.2)-(2.4) is given by the following set of nonlinear equations, the discrete unknowns of which are $U = (U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$:

$$\begin{aligned}\frac{U_K^{n+1}-U_K^n}{\delta t^n}\, m(K) \quad &+ \sum_{L\in\mathcal{N}_K}\Big[(q_{K,L}^{n+1})^+ f(U_K^{n+1}) - (q_{K,L}^{n+1})^- f(U_L^{n+1})\Big] \\ &- \sum_{L\in\mathcal{N}_K}\tau_{K|L}(\varphi(U_L^{n+1})-\varphi(U_K^{n+1})) \\ &- \sum_{\sigma\in\mathcal{E}_{ext,K}}\tau_\sigma(\varphi(\bar{U}_\sigma^{n+1})-\varphi(U_K^{n+1})) \qquad = 0, \\ \forall K\in\mathcal{T}, \forall n\in[\![0,N]\!].\end{aligned} \tag{2.12}$$

where $(q_{K,L}^{n+1})^+$ and $(q_{K,L}^{n+1})^-$ denote the positive and negative parts of $q_{K,L}^{n+1}$ (*i.e.* $(q_{K,L}^{n+1})^+ = \max(q_{K,L}^{n+1},0)$ and $(q_{K,L}^{n+1})^- = -\min(q_{K,L}^{n+1},0)$).

Remark 2.2.2 *The upwind discretization of the flux* $\mathbf{q}f(u)$ *in (2.12) uses the monotonicity of* f *and should be replaced by a flux splitting scheme in the general case.*

Remark 2.2.3 *Thanks to Hypothesis (H6), one gets for all* $K\in\mathcal{T}$ *and* $n\in[\![0,N]\!]$*,*
$\displaystyle\sum_{L\in\mathcal{N}_K} q_{K,L}^{n+1} = \sum_{L\in\mathcal{N}_K}[(q_{K,L}^{n+1})^+ - (q_{K,L}^{n+1})^-] = 0.$ *This leads to*

$$\sum_{L\in\mathcal{N}_K}(q_{K,L}^{n+1})^+ f(U_K^{n+1}) - (q_{K,L}^{n+1})^- f(U_L^{n+1}) = -\sum_{L\in\mathcal{N}_K}(q_{K,L}^{n+1})^-(f(U_L^{n+1})-f(U_K^{n+1})). \tag{2.13}$$

This property will be used in the following.

In Section (2.3) we shall prove the existence (Lemma 2.3.1) and the uniqueness (Lemma 2.3.4) of the solution $U = (U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ to (2.10)-(2.12) . We may then define the approximate solution to (2.2)-(2.4) associated to an admissible discretization $\mathcal{D}$ of $\Omega\times(0,T)$ by:

Definition 2.2.4 *Let* $\mathcal{D}$ *be an admissible discretization of of* $\Omega\times(0,T)$ *in the sense of Definition 2.2.3. The approximate solution of (2.2)-(2.4) associated to the discretization* $\mathcal{D}$ *is defined almost everywhere in* $\Omega\times(0,T)$ *by:*

$$u_{\mathcal{D}}(x,t) = U_K^{n+1},\ \forall x\in K,\ \forall t\in(t^n,t^{n+1}),\ \forall K\in\mathcal{T},\ \forall n\in[\![0,N]\!], \tag{2.14}$$

where $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ *is the unique solution to (2.10)-(2.12) .*

Theorem 2.2.1 (Convergence of the approximate solution towards the entropy weak solution)

Let $\xi\in\mathbb{R}_+^*$*, consider a family of admissible discretizations* $\mathcal{D}$ *of* $\Omega\times(0,T)$ *in the sense of Definition 2.2.3 such that* $\xi\le\mathrm{reg}(\mathcal{D})$*. For a given admissible mesh* $\mathcal{D}$ *of this family, let* $u_{\mathcal{D}}$ *denote the associated approximate solution as defined in Definition 2.2.4. Then:*

$$u_{\mathcal{D}} \longrightarrow u\in L^p(\Omega\times(0,T) \text{ as } \mathrm{size}(\mathcal{D})\longrightarrow 0,\ \forall p\in[1,+\infty),$$

where u *is the unique entropy weak solution to (2.2)-(2.4) .*

The proof of this convergence theorem will be concluded in Section 2.7 after we lay out the properties of the discrete solution (sections 2.3 and 2.4), its convergence towards an ''entropy process solution" (Section 2.5) and a uniqueness result on this entropy process solution (Section 2.6).

Remark 2.2.4 *All the results of this paper also hold for explicit schemes, under a convenient CFL condition on the time step and mesh size.*

2.3 Existence, uniqueness and discrete properties

We state here the properties and estimates which are satisfied by the scheme which we introduced in the previous section and prove existence and uniqueness of the solution to this scheme. All the discrete properties which we address here correspond to natural estimates which are satisfied, at least formally, by regular continuous solutions. Let us first start by an L^∞ estimate:

Lemma 2.3.1 (L^∞ estimate) *Under hypothesis (H), let $\mathcal{D}$ be a discretization of $\Omega \times (0,T)$ in the sense of definition 2.2.3 and let $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ be a solution of scheme (2.10)-(2.12) . Then*

$$U_I \leq U_K^{n+1} \leq U_S, \quad \forall K \in \mathcal{T},\ \forall n \in [\![0,N]\!].$$

Proof.
Let $U_M = \max\limits_{L\in\mathcal{T},m\in[\![0,N]\!]} U_L^{m+1}$ and let $n \in [\![0,N]\!]$ and $K \in \mathcal{T}$ such that $U_K^{n+1} = U_M$. Equations (2.12) and (2.13) yield

$$\begin{aligned} U_M = U_K^{n+1} = U_K^n \quad &+ \quad \frac{\delta t^n}{m(K)} \sum_{L\in\mathcal{N}_K} (q_{K,L}^{n+1})^-(f(U_L^{n+1}) - f(U_K^{n+1})) \\ &+ \quad \frac{\delta t^n}{m(K)} \sum_{L\in\mathcal{N}_K} \tau_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1})) \\ &+ \quad \frac{\delta t^n}{m(K)} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_K^{n+1})). \end{aligned}$$

If one assumes that
$U_M \geq \max\limits_{\sigma\in\mathcal{E}_{ext},m\in[\![0,N]\!]} \bar{U}_\sigma^{m+1}$, using the monotonicity of φ and f, one gets $U_M \leq U_K^n$, and therefore $U_M \leq U_K^0$.
This shows that

$$U_M \leq \max\Big(\max_{\sigma\in\mathcal{E}_{ext},m\in[\![0,N]\!]} \bar{U}_\sigma^{m+1}, \max_{L\in\mathcal{T}} U_L^0\Big),$$

yielding $U_M \leq U_S$. By the same method, one shows that $\min\limits_{L\in\mathcal{T},m\in[\![0,N]\!]} U_L^{m+1} \geq U_I$. □

A corollary of Lemma 2.3.1 is the existence of a solution $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ to (2.10)-(2.12) . (Uniqueness is proven in Lemma (2.3.4) below).

Corollary 2.3.1 (Existence of the solution to the scheme) *Under hypothesis (H), let $\mathcal{D}$ be a discretization of $\Omega \times (0,T)$ in the sense of definition 2.2.3. Then there exists a solution $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ to the scheme (2.10)-(2.12) .*

The proof of this corollary is an adaptation of the technique which was used in [EGGH98] for the existence of the solution to an implicit finite volume scheme for the discretization of a pure hyperbolic equation. The two following lemmata express the monotonicity of the scheme. Both are used to derive continuous entropy inequalities.

Lemma 2.3.2 (Regular convex discrete entropy inequalities) *Under hypothesis (H), let $\mathcal{D}$ be a discretization of $\Omega \times (0,T)$ in the sense of definition 2.2.3 and let $U = (U_K^{n+1})_{K\in\mathcal{T}, n\in[\![0,N]\!]}$ be a solution to (2.10)-(2.12) .*

Then, for all $\eta \in C^2(\mathbb{R},\mathbb{R})$, with $\eta'' \geq 0$, for all μ and ν in $C^1(\mathbb{R},\mathbb{R})$ with $\mu' = \eta'(\varphi)$ and $\nu' = \eta'(\varphi)f'$, for all $K \in \mathcal{T}$, and $n \in [\![0,N]\!]$, there exist $(U_{K,L}^{n+1})_{L\in\mathcal{N}_K}$ with $U_{K,L}^{n+1} \in (\min(U_K^{n+1}, U_L^{n+1}), \max(U_K^{n+1}, U_L^{n+1}))$ for all $L \in \mathcal{N}_K$ and $(U_{K,\sigma}^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}$ with $U_{K,\sigma}^{n+1} \in (\min(U_K^{n+1}, \bar{U}_\sigma^{n+1}), \max(U_K^{n+1}, \bar{U}_\sigma^{n+1}))$ for all $\sigma \in \mathcal{E}_{ext,K}$ satisfying

$$\begin{aligned}
\frac{\mu(U_K^{n+1}) - \mu(U_K^n)}{\delta t^n}\, m(K) \quad &+ \sum_{L\in\mathcal{N}_K} (q_{K,L}^{n+1})^+ \nu(U_K^{n+1}) - (q_{L,K}^{n+1})^- \nu(U_L^{n+1}) \\
&- \sum_{L\in\mathcal{N}_K} \tau_{K|L}(\eta(\varphi(U_L^{n+1})) - \eta(\varphi(U_K^{n+1}))) \\
&- \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma(\eta(\varphi(\bar{U}_\sigma^{n+1})) - \eta(\varphi(U_K^{n+1}))) \\
&+ \frac{1}{2}\sum_{L\in\mathcal{N}_K} \tau_{K|L}\eta''(\varphi(U_{K,L}^{n+1}))(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))^2 \\
&+ \frac{1}{2}\sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma\eta''(\varphi(U_{K,\sigma}^{n+1}))(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_K^{n+1}))^2 \leq 0
\end{aligned} \tag{2.15}$$

Proof.
In order to prove (2.15), one multiplies Equation (2.12) by $\eta'(\varphi(U_K^{n+1}))$.
The convexity of μ yields

$$m(K)\, \frac{U_K^{n+1} - U_K^n}{\delta t^n}\, \eta'(\varphi(U_K^{n+1})) \geq m(K)\, \frac{\mu(U_K^{n+1}) - \mu(U_K^{n+1})}{\delta t^n}.$$

Using the convexity of ν and Remark 2.2.3, one gets

$$\begin{aligned}
-\sum_{L\in\mathcal{N}_K} (q_{K,L}^{n+1})^-(f(U_L^{n+1}) - f(U_K^{n+1}))\eta'(\varphi(U_K^{n+1})) \quad &\geq \quad -\sum_{L\in\mathcal{N}_K} (q_{K,L}^{n+1})^-(\nu(U_L^{n+1}) - \nu(U_K^{n+1})) \\
&\geq \quad \sum_{L\in\mathcal{N}_K} (q_{K,L}^{n+1})^+\nu(U_K^{n+1}) - (q_{K,L}^{n+1})^-\nu(U_L^{n+1})
\end{aligned}$$

The Taylor-Lagrange formula gives, for all $L \in \mathcal{N}_K$ and all $\sigma \in \mathcal{E}_{ext,K}$, the existence of $U_{K,L}^{n+1} \in (\min(U_K^{n+1}, U_L^{n+1}), \max(U_K^{n+1}, U_L^{n+1}))$ and $U_{K,\sigma}^{n+1} \in (\min(U_K^{n+1}, \bar{U}_\sigma^{n+1}), \max(U_K^{n+1}, \bar{U}_\sigma^{n+1}))$ such that

$$(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))\eta'(\varphi(U_K^{n+1})) = (\eta(\varphi(U_L^{n+1})) - \eta(\varphi(U_K^{n+1}))) - \frac{1}{2}\eta''(\varphi(U_{K,L}^{n+1}))(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))^2,$$

$$(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_K^{n+1}))\eta'(\varphi(U_K^{n+1})) = (\eta(\varphi(\bar{U}_\sigma^{n+1})) - \eta(\varphi(U_K^{n+1}))) - \frac{1}{2}\eta''(\varphi(U_{K,\sigma}^{n+1}))(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_K^{n+1}))^2.$$

Then collecting the previous inequalities gives Inequality (2.15) . □

Lemma 2.3.3 (Kruzkov's discrete entropy inequalities) *Under hypothesis* (H)*, let* $\mathcal{D}$ *be a discretization of* $\Omega \times (0,T)$ *in the sense of Definition 2.2.3 and let* $U = (U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ *be a solution of the scheme (2.10)-(2.12) .*

Then, for all $\kappa \in \mathbb{R}$*,* $K \in \mathcal{T}$ *and* $n \in [\![0,N]\!]$*,*

$$\begin{aligned}\frac{|U_K^{n+1}-\kappa| - |U_K^n-\kappa|}{\delta t^n}\, m(K) \quad &+ \sum_{L\in\mathcal{N}_K}\left[\begin{array}{c}(q_{K,L}^{n+1})^+|f(U_K^{n+1})-f(\kappa)| \\ -(q_{K,L}^{n+1})^-|f(U_L^{n+1})-f(\kappa)|\end{array}\right] \\ &- \sum_{L\in\mathcal{N}_K}\tau_{K|L}(|\varphi(U_L^{n+1})-\varphi(\kappa)| - |\varphi(U_K^{n+1})-\varphi(\kappa)|) \\ &- \sum_{\sigma\in\mathcal{E}_{ext,K}}\tau_\sigma(|\varphi(\bar{U}_\sigma^{n+1})-\varphi(\kappa)| - |\varphi(U_K^{n+1})-\varphi(\kappa)|) \le 0\end{aligned} \tag{2.16}$$

Proof. In order to prove Kruzkov's entropy inequalities, one follows [EGGH98]. Equation (2.12) is rewritten as

$$B(U_K^{n+1}, U_K^n, (U_L^{n+1})_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}) = 0,$$

where B is non increasing with respect to each of its arguments except U_K^{n+1}. Consequently,

$$B(U_K^{n+1}, U_K^n\top\kappa, (U_L^{n+1}\top\kappa)_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1}\top\kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \le 0. \tag{2.17}$$

Since $B(\kappa,\kappa,(\kappa)_{L\in\mathcal{N}_K},(\kappa)_{\sigma\in\mathcal{E}_{ext,K}}) = 0$, one gets

$$B(\kappa, U_K^n\top\kappa, (U_L^{n+1}\top\kappa)_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1}\top\kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \le 0. \tag{2.18}$$

Using the fact that $U_K^{n+1}\top\kappa = U_K^{n+1}$ or κ, (2.17) and (2.18) give

$$B(U_K^{n+1}\top\kappa, U_K^n\top\kappa, (U_L^{n+1}\top\kappa)_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1}\top\kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \le 0. \tag{2.19}$$

In the same way one obtains

$$B(U_K^{n+1}\bot\kappa, U_K^n\bot\kappa, (U_L^{n+1}\bot\kappa)_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1}\bot\kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \ge 0. \tag{2.20}$$

Substracting (2.20) from (2.19) and remarking that for any nondecreasing function g and all real values a,b, $g(a\top b) - g(a\bot b) = |g(a)-g(b)|$ yields Inequality (2.16). □
Let us now prove the uniqueness of the solution to (2.10)-(2.12) and define the approximate solution.

Lemma 2.3.4 (Uniqueness of the approximate solution) *Under hypothesis* (H)*, let* $\mathcal{D}$ *be a discretization of* $\Omega\times(0,T)$ *in the sense of Definition 2.2.3. Then there exists a unique solution* $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ *to (2.10)-(2.12) .*

Proof.
The existence of $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ was established in Corollary 2.3.1. There only remains to prove the uniqueness of the solution. Let $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ and $(V_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ (setting $V_K^0 = U_K^0$) be two solutions to the scheme (2.10)-(2.12) . Following the proof of Lemma 2.3.3, one gets, for all $K \in \mathcal{T}$ and all $n \in [\![0, N]\!]$,

$$B(U_K^{n+1}\top V_K^{n+1}, U_K^n \top V_K^n, (U_L^{n+1}\top V_L^{n+1})_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}) \le 0,$$

and

$$B(U_K^{n+1}\perp V_K^{n+1}, U_K^n \perp V_K^n, (U_L^{n+1}\perp V_L^{n+1})_{L\in\mathcal{N}_K}, (\bar{U}_\sigma^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}) \ge 0,$$

which by substraction give

$$\begin{aligned}\frac{|U_K^{n+1}-V_K^{n+1}|-|U_K^n-V_K^n|}{\delta t^n}\, m(K) \quad &+ \sum_{L\in\mathcal{N}_K}\left[\begin{array}{c}(q_{K,L}^{n+1})^+|f(U_K^{n+1})-f(V_K^{n+1})|\\ -(q_{K,L}^{n+1})^-|f(U_L^{n+1})-f(V_L^{n+1})|\end{array}\right]\\ &- \sum_{L\in\mathcal{N}_K}\tau_{K|L}\left[\begin{array}{c}|\varphi(U_L^{n+1})-\varphi(V_L^{n+1})|-\\ |\varphi(U_K^{n+1})-\varphi(V_K^{n+1})|\end{array}\right]\\ &+ \sum_{\sigma\in\mathcal{E}_{ext,K}}\tau_\sigma|\varphi(U_K^{n+1})-\varphi(V_K^{n+1})| \qquad \le 0.\end{aligned} \tag{2.21}$$

For a given $n \in [\![0, N]\!]$, one sums (2.21) on $K \in \mathcal{T}$ and multiplies by δt^n. All the exchange terms between neighboring control volume disappear, and because of the sign of the boundary terms, one gets

$$\sum_{K\in\mathcal{T}}|U_K^{n+1}-V_K^{n+1}|\, m(K) \le \sum_{K\in\mathcal{T}}|U_K^n-V_K^n|\, m(K).$$

Since $U_K^0 = V_K^0$, one concludes $\displaystyle\sum_{K\in\mathcal{T}}|U_K^{n+1}-V_K^{n+1}|\, m(K) = 0$, for all $n \in [\![0, N]\!]$, which concludes the proof of uniqueness. □

Let us now give two discrete estimates on the approximate solution $u_\mathcal{D}$ which will be crucial in the convergence analysis. The first estimate (2.22) is a discrete $L^2(0, T, H^1(\Omega))$ estimate on the function $\zeta(u_\mathcal{D})$ where $\zeta' = \sqrt{\varphi'}$. This estimate will yield some compactness on $\zeta(u_\mathcal{D})$.
The second estimate is the weak BV inequality (2.23) on $f(u_\mathcal{D})$. Formally, it may be seen as the discrete version of a BV estimate for the continuous problem with an additional diffusion term $-\varepsilon\Delta f(u)$. This inequality does not give any compactness property; however it is necessary to prove the convergence of $f(u_\mathcal{D})$ (which is required in order to prove the strong convergence of $u_\mathcal{D}$.) It may be seen as a control of the numerical diffusion which is introduced by the upstream weighting scheme.

Proposition 2.3.1 (Discrete H^1 estimate and weak BV inequality) *Under hypothesis (H), let $\mathcal{D}$ be a discretization of $\Omega \times (0, T)$ in the sense of definition 2.2.3. Let ξ be a real number such that $0 < \xi \le$ $\mathrm{reg}(\mathcal{D})$; let $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ be the solution of the scheme (2.10)-(2.12) .*
Then there exists a real number $C > 0$, only depending on $\Omega, T, u_0, \bar{u}, f, \mathbf{q}, \varphi, \xi$ and $M = \max_{K\in\mathcal{T}} \mathrm{card}\mathcal{E}_K$ such that

$$
\begin{aligned}
(\mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}})))^2 &= \sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_{int}} \tau_{K|L}(\zeta(U_K^{n+1}) - \zeta(U_L^{n+1}))^2 \\
&+ \sum_{n=0}^{N} \delta t^n \sum_{\sigma\in\mathcal{E}_{ext}} \tau_\sigma(\zeta(\bar{U}_\sigma^{n+1}) - \zeta(U_{K_\sigma}^{n+1}))^2 \le C \qquad (2.22)\\
(\mathcal{B}_{\mathcal{D}}(f(u_{\mathcal{D}})))^2 &= \sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_{int}} ((q_{K,L}^{n+1})^- + (q_{K,L}^{n+1})^+)(f(U_K^{n+1}) - f(U_L^{n+1}))^2 \le C \qquad (2.23)
\end{aligned}
$$

Proof. One first defines discrete values by averaging, in each control volume, the function $\bar{u}$, whose trace on $\partial\Omega$ defines the Dirichlet boundary condition. Note that this proof uses $\bar{u} \in H^1(\Omega \times (0,T))$ and not only $\bar{u} \in L^2(0,T;H^1(\Omega))$ and $\bar{u}_t \in L^2(0,T;H^{-1}(\Omega))$, since one uses below the trace of $\bar{u}$ for $t=0$, denoted by $\bar{u}(\cdot,0)$. Let

$$
\bar{U}_K^0 = \frac{1}{m(K)} \int_K \bar{u}(x,0)dx, \qquad \forall K \in \mathcal{T},
$$

$$
\bar{U}_K^{n+1} = \frac{1}{\delta t^n\ m(K)} \int_{t^n}^{t^{n+1}} \int_K \bar{u}(x,t)dxdt, \qquad \forall K \in \mathcal{T}, \forall n \in [\![0,N]\!], \qquad (2.24)
$$

Setting $V = U - \bar{U}$, one multiplies (2.12) by $\delta t^n V_K^{n+1}$ and sums over $K \in \mathcal{T}$ and $n \in [\![0,N]\!]$. This yields $E1 + E2 + E3 = 0$ with

$$
\begin{aligned}
E1 &= \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} m(K)(U_K^{n+1} - U_K^n)V_K^{n+1}, \\
E2 &= \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} ((q_{K,L}^{n+1})^+ f(U_K^{n+1}) - (q_{K,L}^{n+1})^- f(U_L^{n+1}))V_K^{n+1}, \\
E3 &= \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} (\sum_{L\in\mathcal{N}_K} \tau_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))V_K^{n+1} + \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_K^{n+1}))V_K^{n+1}).
\end{aligned}
$$

Using $U = V + \bar{U}$ yields $E1 = E11 + E12$ with

$$
E11 = \frac{1}{2}\sum_{K\in\mathcal{T}} m(K)((V_K^{N+1})^2 - (V_K^0)^2) + \frac{1}{2}\sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)(V_K^{n+1} - V_K^n)^2
$$

$$
E12 = \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)(\bar{U}_K^{n+1} - \bar{U}_K^n)V_K^{n+1}.
$$

One has

$$
E12 = \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)(\bar{U}_K^{n+1} - \frac{1}{m(K)}\int_K \bar{u}(x,t^n)dx)V_K^{n+1} + \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)(\frac{1}{m(K)}\int_K \bar{u}(x,t^n)dx - \bar{U}_K^n)V_K^{n+1}.
$$

and therefore:

$$E12 \leq 2[(\sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K))A_{n,K}V_K^{n+1})^2 + (\sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)B_{n,K}V_K^{n+1})^2] \tag{2.25}$$

with

$$A_{n,K} = \bar{U}_K^{n+1} - \frac{1}{m(K)}\int_K \bar{u}(x,t^n)dx \text{ and } B_{n,K} = \frac{1}{m(K)}\int_K \bar{u}(x,t^n) - \bar{U}_K^n.$$

By density one has:

$$|A_{n,K}| \leq \frac{1}{m(K)}\|\bar{u}_t\|_{L^1(K\times(t_n,t_{n+1}))} \text{ and } |B_{n,K}| \leq \frac{1}{m(K)}\|\bar{u}_t\|_{L^1(K\times(t_{n-1},t_n))}.$$

Using these two inequalities and the L^∞ stability of the scheme (Lemma 2.3.1) in (2.25) yields:

$$E12 \leq \|\bar{u}_t\|_{L^1(\Omega\times(0,T))}(\max(U_S,-U_I) + \|\bar{u}\|_{L^\infty(\Omega\times(0,T))}).$$

Now remarking that

$$E11 \geq -\frac{1}{2}\sum_{K\in\mathcal{T}} m(K){V_K^0}^2 \geq -\frac{1}{2}\|u_0 - \bar{u}(\cdot,0)\|^2_{L^2(\Omega)}$$

the previous inequality allows us to obtain the existence of $C1 > 0$, only depending on $\Omega, T, u_0, \bar{u}$ and ξ, such that $E1 \geq C1$.

The term $E2$ can be decomposed in $E2 = E21 + E22$ with

$$\begin{aligned} E21 &= \sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}((q_{K,L}^{n+1})^+ f(U_K^n) - (q_{K,L}^{n+1})^- f(U_L^{n+1}))U_K^{n+1}, \\ E22 &= -\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}((q_{K,L}^{n+1})^+ f(U_K^n) - (q_{K,L}^{n+1})^- f(U_L^{n+1}))\bar{U}_K^{n+1}, \end{aligned}$$

Using Remark 2.2.3, one gets

$$E21 = \sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}(q_{K,L}^{n+1})^-(f(U_K^{n+1}) - f(U_L^{n+1}))U_K^{n+1}. \tag{2.26}$$

Let g be a primitive of $xf'(x)$. The following inequality holds for all pairs of real values (a,b) (see [EGH00b]).

$$g(b) - g(a) \leq b(f(b) - f(a)) - \frac{1}{2F}(f(b) - f(a))^2 \tag{2.27}$$

Using (2.27) for $(a,b) = (U_L^{n+1}, U_K^{n+1})$ and (2.26) yield

$$E21 \geq \sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}(q_{K,L}^{n+1})^-(g(U_K^{n+1}) - g(U_L^{n+1})) + \frac{1}{2F}(\mathcal{B}_{\mathcal{D}}(f(U)))^2.$$

Using Remark 2.2.3 with g instead of f gives

$$\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} (q_{K,L}^{n+1})^-(g(U_K^{n+1}) - g(U_L^{n+1})) \quad = \quad 0,$$

and therefore

$$E21 \geq \frac{1}{2F}(\mathcal{B}_{\mathcal{D}}(f(U)))^2.$$

A discrete space integration by parts in $E22$ does not yield any boundary term since $\mathbf{q}\cdot\mathbf{n} = 0$ on $\partial\Omega$, and gives, using the Cauchy-Schwarz inequality,

$$\begin{aligned}
E22 &= -\sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_{int}} ((q_{K,L}^{n+1})^+ f(U_K^{n+1}) - (q_{K,L}^{n+1})^- f(U_L^{n+1}))(\bar{U}_K^{n+1} - \bar{U}_L^{n+1}) \\
&\geq -\|\mathbf{q}\|_{L^\infty(\Omega\times(0,T))} \max_{s\in[U_I,U_S]} |f(s)| \sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_{int}} m(K|L)|\bar{U}_K^{n+1} - \bar{U}_L^{n+1}| \\
&\geq -\|\mathbf{q}\|_{L^\infty(\Omega\times(0,T))} \max_{s\in[U_I,U_S]} |f(s)| \mathcal{N}_{\mathcal{D}}(\bar{U}) [\sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_{int}} m(K|L) d(x_K, x_L)]^{\frac{1}{2}} \\
&\geq -\mathcal{N}_{\mathcal{D}}(\bar{U}) \|\mathbf{q}\|_{L^\infty(\Omega\times(0,T))} \max_{s\in[U_I,U_S]} |f(s)| (d\ m(\Omega)\ T)^{\frac{1}{2}}.
\end{aligned}$$

The following estimate for $\mathcal{N}_{\mathcal{D}}(\bar{U})$ holds (see [EGH99]):

$$\mathcal{N}_{\mathcal{D}}(\bar{U}) \leq F(\xi, M)\|\bar{u}\|_{L^2(0,T,H^1(\Omega))}, \tag{2.28}$$

where $F \geq 0$ only depends on $\mathrm{reg}(\mathcal{T})$ and $M = \max_{K\in\mathcal{T}} \mathrm{card}\mathcal{E}_K$ leading to a lower bound of $E22$ denoted by $C22$, only depending on $\Omega, T, u_0, \bar{u}, f, \mathbf{q}, \xi$ and M.

There only remains to deal with $E3$. A discrete space integration by parts, using the fact that $V_\sigma^{n+1} = 0, \forall \sigma \in \mathcal{E}_{ext}, \forall n \in [\![0, N]\!]$, yields

$$\begin{aligned}
E3 = \sum_{n=0}^{N} \delta t^n (& \sum_{K|L\in\mathcal{E}_{int}} \tau_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))(V_L^{n+1} - V_K^{n+1}) \\
& + \sum_{\sigma\in\mathcal{E}_{ext}} \tau_\sigma(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_{K_\sigma}^{n+1}))(V_\sigma^{n+1} - V_{K_\sigma}^{n+1}).
\end{aligned}$$

Writing again V into $U - \bar{U}$ leads to $E3 = E31 + E32$ where

$$\begin{aligned}
E31 = \sum_{n=0}^{N} \delta t^n (& \sum_{K|L\in\mathcal{E}_{int}} \tau_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))(U_L^{n+1} - U_K^{n+1}) \\
& + \sum_{\sigma\in\mathcal{E}_{ext}} \tau_\sigma(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_{K_\sigma}^{n+1}))(\bar{U}_\sigma^{n+1} - U_{K_\sigma}^{n+1})
\end{aligned}$$

$$
\begin{aligned}
E32 \quad = \quad & -\sum_{n=0}^{N} \delta t^n (\sum_{K|L\in\mathcal{E}_{int}} \tau_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))(\bar{U}_L^{n+1} - \bar{U}_K^{n+1}) \\
& + \sum_{\sigma\in\mathcal{E}_{ext}} \tau_\sigma(\varphi(\bar{U}_\sigma^{n+1}) - \varphi(U_{K_\sigma}^{n+1}))(\bar{U}_\sigma^{n+1} - \bar{U}_{K_\sigma}^{n+1})
\end{aligned}
$$

One has for all pairs of real numbers (a,b) the inequality $(\zeta(a)-\zeta(b))^2 \leq (a-b)(\varphi(a)-\varphi(b))$. Also using $\varphi' \leq \sqrt{\Phi}\zeta'$ (recall that $\Phi = \|\varphi'\|_\infty$), one gets

$$
\begin{aligned}
E31 \quad & \geq \quad (\mathcal{N}_\mathcal{D}(\zeta(U)))^2, \\
E32 \quad & \geq \quad -\sqrt{\Phi}\mathcal{N}_\mathcal{D}(\zeta(U))\mathcal{N}_\mathcal{D}(\bar{U}).
\end{aligned}
$$

Using the Young inequality and (2.28), one gets the existence of $C32$ only depending on $\Omega, T, u_0, \bar{u}, f, \mathbf{q}, \varphi, \xi$ such that

$$
E32 \quad \geq \quad -\frac{1}{2}(\mathcal{N}_\mathcal{D}(\zeta(U)))^2 + C32.
$$

Gathering the previous inequalities, one gets

$$
C1 + \frac{1}{2F}(\mathcal{B}_\mathcal{D}(f(U)))^2 + C22 + \frac{1}{2}(\mathcal{N}_\mathcal{D}(\zeta(U)))^2 + C32 \leq 0,
$$

which completes the proof. □

Remarking that from the estimate of Lemma 2 in [EGH99], one has $\mathcal{N}_\mathcal{D}(\zeta(\bar{U})) \leq \sqrt{\Phi}C\|\bar{u}\|_{L^2(0,T,H^1(\Omega))}$, where $C \geq 0$ only depends on $\mathrm{reg}(\mathcal{T})$ and $M = \max_{K\in\mathcal{T}} \mathrm{card}(\mathcal{E}_K)$, one gets

Corollary 2.3.2 (Discrete H_0^1 estimate) *Under hypothesis (H), let $\mathcal{D}$ be a discretization of $\Omega\times(0,T)$ in the sense of definition 2.2.3. Let ξ be a real number such that $0 < \xi \leq \mathrm{reg}(\mathcal{D})$, let $U = (U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ be the solution of the scheme (2.10)-(2.12) and let $\bar{U} = (\bar{U}_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ be defined by (2.24). Then, setting $Z = \zeta(U)-\zeta(\bar{U})$, there exists $C' \in \mathbb{R}_+$, only depending on $\Omega, T, u_0, \bar{u}, \varphi, \mathbf{q}, f, \xi$ and $M = \max_{K\in\mathcal{T}} \mathrm{card}(\mathcal{E}_K)$ such that*

$$
\sum_{n=0}^{N} \delta t^n (\sum_{K|L\in\mathcal{E}_{int}} \tau_{K|L}(Z_K^{n+1} - Z_L^{n+1})^2 + \sum_{\sigma\in\mathcal{E}_{ext}} \tau_\sigma(Z_{K_\sigma}^{n+1})^2) \leq C'
$$

2.4 Compactness of a family of approximate solutions

From Lemma 2.3.1, we know that for any sequence of admissible discretizations $(\mathcal{D}_m)_{m\in\mathbb{N}}$, $\Omega\times(0,T)$ in the sense of Definition 2.2.3, the associated sequence of approximate solutions $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ is bounded in $L^\infty(\Omega\times(0,T))$. Therefore one may extract a subsequence which converges for the weak star topology of $L^\infty(\Omega\times(0,T))$ as m tends to infinity. This convergence is unfortunately insufficient to pass to the limit in the nonlinearities. In order to pass to the limit, we shall use two tools:

1. the nonlinear weak star convergence which was introduced in [EGGH98] and which is equivalent to the notion of convergence towards a Young measure as developed in [DiP85].

2. Kolmogorov's compactness theorem, which was used in [EGH99] in the case of a semi-linear elliptic equation.

Theorem 2.4.1 (Nonlinear weak star convergence) *Let Q be a Borel subset of $\mathbb{R}^k$ and $(u_n)_{n\in\mathbb{N}}$ be a bounded sequence in $L^\infty(Q)$. Then there exists $u \in L^\infty(Q\times(0,1))$, such that up to a subsequence, u_n tends to u "in the nonlinear weak sense" as $n \longrightarrow \infty$, i.e.:*

$$\forall g \in \mathcal{C}(\mathbb{R},\mathbb{R}),\ g(u_n) \rightharpoonup \int_0^1 g(u(\cdot,\alpha))d\alpha \text{ for the weak star topology of } L^\infty(Q) \text{ as } n \longrightarrow \infty.$$

We refer to [DiP85, EGGH98] for details and proof of Theorem 2.4.1.

This compactness result allows us to exhibit a limit (in the nonlinear weak sense) $u \in L^\infty(\Omega\times(0,T)\times(0,1))$ of a subsequence of the sequence $u_{\mathcal{D}_m}$ which we considered above. Of course, in order to show that this function u is the unique entropy weak solution to (2.2)-(2.4) , we shall need to show that it does not depend on its variable α and that is satisfies the boundary condition (2.7) and the entropy inequalities (2.8) of Definition 2.1.1.

Let us now turn to the the Riesz-Fréchet-Kolmogorov compactness criterion (see e.g. [Bré83]) which will allow us to pass to the limit in the nonlinear second order terms.

Theorem 2.4.2 (Riesz-Fréchet-Kolmogorov) *Let Q be an open bounded subset of $\mathbb{R}^k$ and $(u_n)_{n\in\mathbb{N}}$ be a bounded sequence in $L^2(\mathbb{R}^k)$ such that*

$$\lim_{|\delta|\to 0}[\sup_{n\in\mathbb{N}} \|u_n(\cdot+\delta)-u_n(\cdot)\|_{L^2(Q)}] = 0,$$

then there exists $u \in L^2(Q)$ such that, up to a subsequence,

$$u_n \to u \text{ in } L^2(Q) \text{ as } n \longrightarrow \infty. \text{ as } n \longrightarrow \infty.$$

Let us now show that we are in position to apply the Riesz-Fréchet-Kolmogorov to $(\zeta(u_{\mathcal{D}_m}))_{m\in\mathbb{N}}$. From the discrete estimates Proposition 2.3.1 and Corollary 2.3.2, one can state the following continuous estimates and $z_{\mathcal{D}}$, where $z_{\mathcal{D}}$ is defined almost everywhere in $\Omega\times(0,T)$ by

$$z_{\mathcal{D}}(x,t) = \zeta(U_K^{n+1}) - \zeta(\bar{U}_K^{n+1}) \text{ for } x \in K \text{ and } t \in (t^n, t^{n+1}) \tag{2.29}$$

where $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ is the solution to (2.10)-(2.12) and $(\bar{U}_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ is defined by (2.24).

Corollary 2.4.1 (Space and time translates estimates) *Under hypothesis (H), let $\mathcal{D}$ be a discretization of $\Omega\times(0,T)$ in the sense of Definition 2.2.3. Let ξ be a real number such that $0 < \xi \leq \mathrm{reg}(\mathcal{D})$; let U be the solution of scheme (2.10)-(2.12) , and let $u_{\mathcal{D}}$ be defined by (2.14). Let $\bar{U}$ be defined by (2.24), let $z_{\mathcal{D}}$ be defined by (2.29), and be prolonged by zero on $(0,T)\times\Omega^c$. Then there exist C_1 only depending on $\Omega, T, u_0, \bar{u}, \varphi, \mathbf{q}, f, \xi$ and $M = \max_{K\in\mathcal{T}} \mathrm{card}(\mathcal{E}_K)$, and C_0, only depending on Ω, such that*

$$\forall \xi \in \mathbb{R}^d, \int_0^T \int_{\mathbb{R}^d} (z_{\mathcal{D}}(x+\xi,t) - z_{\mathcal{D}}(x,t))^2 dxdt \leq C_1|\xi|(|\xi| + C_0\,\mathrm{size}(\mathcal{T})), \tag{2.30}$$

and there exists C_2 only depending on $\Omega, T, u_0, \bar{u}, \varphi, \mathbf{q}, f, \xi$ and M such that

$$\forall s > 0, \int_0^{T-s} \int_{\mathbb{R}^d} (\zeta(u_{\mathcal{D}})(x,t+s) - \zeta(u_{\mathcal{D}})(x,t))^2 dxdt \le C_2\ s. \tag{2.31}$$

The use of space translate estimates for the study of numerical schemes for elliptic problems was recently introduced in [EGH99]. The technique of [EGH99] may easily be adapted here to prove (2.30), using the estimates of Corollary 2.3.2. A time translate estimate was introduced in [EGHNS98] to obtain some compactness in the study of finite volume schemes for parabolic equations. The proof of (2.31) follows the technique of [EGHNS98] and uses estimate (2.22) and the discrete equation (2.12).

From Theorem 2.4.2 and the estimates (2.30) and (2.31) of Corollary 2.4.1 we deduce the following compactness result:

Corollary 2.4.2 (Compactness of a family of approximate solutions) *Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of discretizations of $\Omega\times(0,T)$ in the sense of definition 2.2.3 such that there exists $\xi > 0$ with* $\mathrm{reg}(\mathcal{D}_m) \ge \xi$ *for all $m \in \mathbb{N}$. For all $m \in \mathbb{N}$, let $u_{\mathcal{D}_m}$ be defined by the scheme (2.10)-(2.12) and (2.14) with $\mathcal{D} = \mathcal{D}_m$, and let $z_{\mathcal{D}_m}$ e defined by (2.29) with $\mathcal{D} = \mathcal{D}_m$ and (2.24). Then there exists $u \in L^\infty(\Omega\times(0,T)\times(0,1))$ and $z \in L^2(\Omega\times(0,T))$ such that, up to a subsequence, $u_{\mathcal{D}_m} \rightharpoonup u$ in the nonlinear weak star sense and $z_{\mathcal{D}_m} \to z$ in $L^2(\Omega\times(0,T)$ as $m\to\infty$. Furthermore, $z \in L^2(0,T,H_0^1(\Omega))$, $\zeta(u) = z - \zeta(\bar u)$, $\zeta(u_{\mathcal{D}_m})$ converges to $\zeta(u)$ in $L^2(0,T,H^1(\Omega)$ and $\zeta(u) = \zeta(\bar u)$ a.e. on $\partial\Omega$.*

Proof. The convergence of $u_{\mathcal{D}_m}$ towards $u \in L^\infty(\Omega\times(0,T)\times(0,1))$ in the nonlinear weak star sense is a consequence of Lemma 2.3.1 and Theorem 2.4.1. The convergence of $z_{\mathcal{D}_m}$ to z in $L^2(\Omega\times(0,T)$ is a consequence of Theorem 2.4.2 and the estimates (2.30) and (2.31) of Corollary 2.4.1.

Following [EGH00b] or [EGH99], one then deduces from (2.31) that $D_i z \in L^2(\Omega\times(0,T)$ for $i = 1,\dots,d$ and since $z_{\mathcal{D}_m}(x,t) = 0$ on $\Omega^c\times(0,T)$ for all $m\in\mathbb{N}$, one has $z \in L^2(0,T,H_0^1(\Omega))$.
Now since $u_{\mathcal{D}_m}$ converges to u in the nonlinear weak star sense and that the function $\bar u_{\mathcal{D}_m}$ defined a.e. by $\bar u_{\mathcal{D}_m}(x,t) = \bar U_K^{n+1}$ for (x,t) in $K\times(t^n,t^{n+1})$ converges uniformly to $\bar u$, one deduces that $\zeta(u_{\mathcal{D}_m})$ converges to $\zeta(u)$ in the nonlinear weak star sense and to $z+\zeta(\bar u)$ in $L^2(\Omega\times(0,T)$ as m tends to infinity. Therefore, by Lemma 2.4.1 below, one obtains that $\zeta(u) = z+\zeta(\bar u)$ and $\zeta(u)$ does not depend on α. Furthermore, since $z\in L^2(0,T,H_0^1(\Omega))$, it follows that $\zeta(u) = \zeta(\bar u)$ a.e. on $\partial\Omega$ which ends the proof of the corollary.
□

Lemma 2.4.1 *Let Q be a Borel subset of $\mathbb{R}^k$ and let $(u_n)_{n\in\mathbb{N}} \subset L^\infty(Q)$ be such that u_n converges to $u \in L^\infty(Q\times(0,1))$ in the nonlinear weak star sense, and to w in $L^2(Q)$, as n tends to infinity, then $u(x,\alpha) = w(x)$, for a.e. $(x,\alpha)\in Q\times(0,1)$ and u does not depend on α.*

Proof. With the notations of the lemma, we have

$$\int_0^1\int_Q (u(x,\alpha)-w(x))^2 dxd\alpha = \int_0^1\int_Q (u(x,\alpha))^2 dxd\alpha - 2\int_0^1\int_Q u(x,\alpha)w(x)dxd\alpha + \int_0^1\int_Q w(x)^2 dxd\alpha.$$

Since u_n tends to u in the nonlinear weak star sense , one has

$$\int_0^1\int_Q (u(x,\alpha))^2 dxd\alpha = \lim_{n\longrightarrow+\infty}\int_Q (u_n(x))^2dx \ \text{ and }\ \int_0^1\int_Q u(x,\alpha)w(x)dxd\alpha = \lim_{n\longrightarrow+\infty}\int_Q u_n(x)w(x)dx,$$

and since u_n tends to w in $L^2(Q)$, one deduces that $u(x,\alpha) = w(x)$, for a.e. $(x,\alpha)\in Q\times(0,1)$ and u does not depend on α. □

2.5 Convergence towards an entropy process solution

This section is mainly devoted to the proof of the convergence theorem 2.5.1, which states the convergence of the approximate solution to a measure valued solution as introduced in [DiP85], which is also called entropy process solution [EGGH98], and defined as follows.

Definition 2.5.1 *Under hypothesis (H), an entropy process solution to Problem (2.2)-(2.4) is a function u such that,*

$$u \in L^\infty(\Omega \times (0,T) \times (0,1)),$$

$$\varphi(u) - \varphi(\bar{u}) \in L^2(0,T; H_0^1(\Omega)),$$

(note that $\varphi(u)$ does not depend on α, and u satisfies the following inequalities:

1. Regular convex entropy inequalities:

$$\int_{\Omega\times(0,T)} \begin{bmatrix} \int_0^1 \mu(u(x,t,\alpha))d\alpha\ \psi_t(x,t)+ \\ \int_0^1 \nu(u(x,t,\alpha))d\alpha\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t) \\ -\nabla\eta(\varphi(u)(x,t))\cdot\nabla\psi(x,t) \\ -\eta''(\varphi(u)(x,t))(\nabla\varphi(u)(x,t))^2\psi(x,t) \end{bmatrix} dxdt + \int_\Omega \mu(u_0(x))\psi(x,0)dx \geq 0, \tag{2.32}$$

$$\forall\ \psi \in \mathcal{D}^+(\Omega\times[0,T)),\ \forall\eta \in C^2(\mathbb{R}), \eta'' \geq 0,\ \mu' = \eta'(\varphi(\cdot)),\ \nu' = \eta'(\varphi(\cdot))f'(\cdot).$$

2. Kruzkov's entropy inequalities:

$$\int_{\Omega\times(0,T)} \begin{bmatrix} \int_0^1 |u(x,t,\alpha)-\kappa| d\alpha\ \psi_t(x,t)+ \\ \int_0^1 (f(u(x,t,\alpha)\top\kappa) - f(u(x,t,\alpha)\bot\kappa))d\alpha\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t) \\ -\nabla|\varphi(u)(x,t)-\varphi(\kappa)|\cdot\nabla\psi(x,t) \end{bmatrix} dxdt + \int_\Omega |u_0(x)-\kappa|\psi(x,0)dx \geq 0,$$

$$\forall\ \psi \in \mathcal{D}^+(\Omega\times[0,T)),\ \forall\kappa \in \mathbb{R}. \tag{2.33}$$

This notion of entropy process solution appears to be the natural limit of the approximate solutions. This is expressed in the following theorem.

Theorem 2.5.1 (Convergence towards an entropy process solution) *Under hypothesis (H), let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of discretizations of $\Omega\times(0,T)$ in the sense of Definition 2.2.3, with $\mathrm{size}(\mathcal{D}_m)\to 0$ as $m\to\infty$, such that there exists $\xi > 0$ with $\mathrm{reg}(\mathcal{D}_m) \geq \xi$ for all $m \in \mathbb{N}$. For all $m \in \mathbb{N}$, let $u_{\mathcal{D}_m}$ be defined by the scheme (2.10)-(2.12) and (2.14) with $\mathcal{D} = \mathcal{D}_m$.*
Then, there exists an entropy process solution of (2.2)-(2.4) in the sense of Definition 2.5.1 and a subsequence of $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$, again denoted by $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$, such that $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ converges to u in the nonlinear weak star sense and $(\zeta(u_{\mathcal{D}_m}))_{m\in\mathbb{N}}$ converges in $L^2(\Omega\times(0,T))$ to $\zeta(u) \in L^2(0,T;H^1(\Omega))$ as m tends to ∞.

Proof. By Lemma 2.4.2, there exist $u \in L^\infty(\Omega\times(0,T)\times(0,1))$ and a subsequence of $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$, again denoted $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$, such that $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ converges to u in the nonlinear weak star topology and $(\zeta(u_{\mathcal{D}_m}))_{m\in\mathbb{N}}$ converges in $L^2(\Omega\times(0,T))$ to $\zeta(u) \in L^2(0,T;H^1(\Omega))$. There remains to show that the function $u \in L^\infty(\Omega\times(0,T)\times(0,1))$ is an entropy process solution.

A number of the arguments involved in order to do so may be found in [EGGH98] or [EGHNS98] and therefore will be given with few details. The main new argument introduced here concerns the term $\int_{\Omega\times(0,T)} \eta''(\varphi(u)(x,t))(\nabla\varphi(u)(x,t))^2\psi(x,t)dxdt$ in equation (2.32). The passage to the limit to obtain this nonlinearity motivates the use of the technical lemma 2.5.2 below (a related technique was used in [GHM99] in the case of a variational inequality).

The idea of the proof is to derive the continuous inequalities (2.32) and (2.33) for the limit u by mutliplying the discrete entropy inequalities (2.15) and (2.16) by regular test functions and passing to the limit. Indeed, let $\psi \in \mathcal{D}^+(\Omega\times[0,T)) = \{\psi \in C_c^\infty(\Omega\times\mathbb{R},\mathbb{R}_+), \psi(\cdot,T)=0\}$. For a given m, let us denote $\mathcal{D}=\mathcal{D}_m$, and let $(U_K^{n+1})_{K\in\mathcal{T},n\in[\![0,N]\!]}$ be the solution of the scheme (2.10)-(2.12) associated to $\mathcal{D}$. Let $\Psi = (\Psi_K^n)_{K\in\mathcal{T},n\in[\![0,N+1]\!]}$ be defined by

$$\Psi_K^n = \psi(x_K,t^n) \qquad \forall K\in\mathcal{T}, \forall n\in[\![0,N+1]\!].$$

Remark 2.5.1 *One cannot use for Ψ_K^n the mean value of ψ on $K\times(t^n,t^{n+1})$; indeed, in order to pass to the limit on the term $A3_\mathcal{D}$ below (see (2.34) and (2.35)), we shall use the consistency of the approximation $\frac{\Psi_K^n)-\Psi_L^n}{d(x_K,x_L)}$ to the normal derivative $\nabla\psi\cdot\mathbf{n}_{K,L}$. This consistency holds if $\Psi_K^n=\psi(x_K,t^n)$ thanks to the assumption on the family $(x_K)_{K\in\mathcal{T}}$ in Definition 2.2.3, but does not generally hold if Ψ_K^n is the mean value of ψ on $K\times(t^n,t^{n+1})$.*
Note that discrete values using the mean values were used for $\bar{u}$ when studying an upper bound of $\mathcal{N}_\mathcal{D}(\bar{U})$ with respect to the $L^2(0,T;H^1(\Omega))$ norm of $\bar{u}$. However we did not have to use the consistency of the flux on $\bar{u}$.

With the notations of lemmas 2.3.2 and 2.3.3, let us multiply the discrete entropy inequalities (2.15) and (2.16) by $\delta t^n\Psi_K^n$ and sum over $K\in\mathcal{T}$ and $n\in[\![0,N]\!]$. From (2.15), one gets $A1_\mathcal{D}+A2_\mathcal{D}+A3_\mathcal{D}+A4_\mathcal{D}\le 0$ with

$$\begin{aligned}
A1_\mathcal{D} &= \sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}} \quad m(K)\frac{\mu(U_K^{n+1})-\mu(U_K^n)}{\delta t^n}\Psi_K^n\\
A2_\mathcal{D} &= -\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}} \quad \sum_{L\in\mathcal{N}_K}((q_{K,L}^{n+1})^-(\nu(U_L^{n+1})-\nu(U_K^{n+1})))\Psi_K^n\\
A3_\mathcal{D} &= -\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}} \quad (\sum_{L\in\mathcal{N}_K}\tau_{K|L}(\eta(\varphi(U_L^{n+1}))-\eta(\varphi(U_K^{n+1})))\Psi_K^n\\
&\qquad +\sum_{\sigma\in\mathcal{E}_{ext,K}}\tau_\sigma(\eta(\varphi(\bar{U}_\sigma^{n+1})-\eta(\varphi(U_K^{n+1})))\Psi_K^n)\\
A4_\mathcal{D} &= \sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}} \quad (\frac{1}{2}\sum_{L\in\mathcal{N}_K}\tau_{K|L}\eta''(\varphi(U_{K,L}^{n+1}))(\varphi(U_L^{n+1})-\varphi(U_K^{n+1}))^2\Psi_K^n\\
&\qquad +\frac{1}{2}\sum_{\sigma\in\mathcal{E}_{ext,K}}\tau_\sigma\eta''(\varphi(U_{K,\sigma}^{n+1}))(\varphi(\bar{U}_\sigma^{n+1})-\varphi(U_K^{n+1}))^2\Psi_K^n)
\end{aligned}$$

Each of these terms will be shown to converge to the corresponding continuous terms of Inequality (2.32) by passing to the limit on the space and time steps, i.e. letting $m\to\infty$.

Since $\psi(\cdot,T)=0$, one has $\Psi_K^{N+1}=0$ and therefore:

$$\begin{aligned} A1_{\mathcal{D}} &= \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} m(K)\mu(U_K^{n+1}) \frac{\Psi_K^n - \Psi_K^{n+1}}{\delta t^n} \\ &\quad - \sum_{K\in\mathcal{T}} m(K)\Psi_K^0 \mu(u_K^0) \end{aligned}$$

The sequence $\mu(u_{\mathcal{D}})$ converges weakly to $\int_0^1 \mu(u(\cdot,\alpha))d\alpha$ as $m\to\infty$. Let $\chi_{\mathcal{D}}$ be the function defined almost everywhere on $\Omega\times(0,T)$ by $\chi_{\mathcal{D}}(x,t) = \frac{\Psi_K^n - \Psi_K^{n+1}}{\delta t^n}$ if $(x,t)\in K\times(t^n,t^{n+1})$; then $\chi_{\mathcal{D}}$ converges to ψ_t in $L^1(\Omega\times(0,T))$ as $m\longrightarrow+\infty$. Furthermore, let $\psi_{\mathcal{T}}^0$ (resp $u_{\mathcal{T}}^0$) be defined almost everywhere on Ω by $\psi_{\mathcal{T}}^0 = \Psi_K^0$ (resp. $u_{\mathcal{T}}^0 = U_K^0$) if $x\in K$. Then, $\mu(u_{\mathcal{T}}^0)$ converges to $\mu(u_0)$ in $L^p(\Omega)$ for any $p\in[1,+\infty)$ and $\psi_{\mathcal{T}}^0$ converges to $\psi(.,0)$ uniformly as $m\longrightarrow+\infty$. Hence passing to the limit as $m\longrightarrow+\infty$ in $A1_{\mathcal{D}}$ yields:

$$\lim_{m\to\infty} A1_{\mathcal{D}_m} = -\int_0^T\int_\Omega\int_0^1 \mu(u(x,t,\alpha))d\alpha\,\psi_t(x,t)dxdt - \int_0^T\int_\Omega \mu(u_0(x))\psi(x,0)dx.$$

Let us now rewrite $A2_{\mathcal{D}}$ as:

$$A2_{\mathcal{D}} = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} \nu(U_K^{n+1})((q_{K,L}^{n+1})^+\Psi_L^n - (q_{K,L}^{n+1})^-\Psi_K^n).$$

We replace the term $(q_{K,L}^{n+1})^+\Psi_L^n - (q_{K,L}^{n+1})^-\Psi_K^n$ by $\frac{1}{\delta t^n}\int_{t^n}^{t^{n+1}}\int_{K|L}\psi(x,t)\,\mathbf{q}(x,t)\cdot\mathbf{n}_{K,L}d\gamma(x)dt$. When doing so, we commit an error which may be controlled (see the details in [EGGH98]) thanks to the consistency and the conservativity of the scheme and thanks to the the weak BV inequality (2.23). Using the weak convergence of $\nu(u_{\mathcal{T}})$ to $\int_0^1 \nu(u(\cdot,\alpha))d\alpha$ as $m\to\infty$, we then obtain:

$$\begin{aligned} \lim_{m\to\infty} A2_{\mathcal{D}_m} &= -\int_0^T\int_\Omega\int_0^1 \nu(u(x,t,\alpha))d\alpha\nabla(\mathbf{q}(x,t)\,\psi(x,t))dxdt \\ &= -\int_0^T\int_\Omega\int_0^1 \nu(u(x,t,\alpha))d\alpha\,\mathbf{q}(x,t)\cdot\nabla\psi(x,t)dxdt. \end{aligned}$$

Turning now to the study of $A3_{\mathcal{D}}$, one remarks that for size$(\mathcal{T})$ small enough, the support of ψ does not intersect the control volumes with edges on $\partial\Omega$. Then for all control volumes $K\in\mathcal{T}$ the sum over $\sigma\in\mathcal{E}_{ext,K}$ vanishes and thus

$$A3_{\mathcal{D}} = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} \tau_{K|L}\eta(\varphi(U_K^n))(\Psi_L^n - \Psi_K^n) \tag{2.34}$$

Using the consistency of $\tau_{K|L}(\Psi_L^n - \Psi_K^n)$ with $\frac{1}{\delta t^n}\int_{t^n}^{t^{n+1}}\int_{K|L}\nabla\psi(x,t)\cdot\mathbf{n}_{K,L}d\gamma(x)dt$, Estimate (2.22) and the (strong) convergence of $\eta(\varphi(u_{\mathcal{D}}))$ to $\eta(\varphi(u))$ as $m\to\infty$, one gets with computations similar as in [EGH99]:

$$\lim_{m\to\infty} A3_{\mathcal{D}_m} = -\int_0^T\int_\Omega \eta(\varphi(u))(x,t)\Delta\psi(x,t)dxdt = \int_0^T\int_\Omega \nabla\eta(\varphi(u))(x,t)\cdot\nabla\psi(x,t)dxdt. \tag{2.35}$$

One now deals with $A4_{\mathcal{D}}$. The second term of $A4_{\mathcal{D}}$ vanishes if size$(\mathcal{T})$ is again sufficiently small. Then $A4_{\mathcal{D}}$ reduces to its first term which writes, after gathering by edges:

$$A4_{\mathcal{D}} \quad = \quad -\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_{int}} \tau_{K|L}\frac{\eta''(\varphi(U_{K,L}^{n+1}))\Psi_K^n+\eta''(\varphi(U_{L,K}^{n+1}))\Psi_L^n}{2}(\varphi(U_L^{n+1})-\varphi(U_K^{n+1}))^2 \quad (2.36)$$

Let us now introduce the sets $\mathcal{V}_\sigma$ for $\sigma\in\mathcal{E}$. Let K be a control volume and $\sigma\in\mathcal{E}_K$. One defines $\mathcal{V}_{K,\sigma}=\{tx_K+(1-t)x, x\in\sigma, t\in(0,1)\}$. For $\sigma=K|L$, $\mathcal{V}_\sigma=\mathcal{V}_{K,\sigma}\cup\mathcal{V}_{L,\sigma}$ and for $\sigma\in\mathcal{E}_{ext,K}$, $\mathcal{V}_\sigma=\mathcal{V}_{K,\sigma}$. One denotes by $H_{K|L}^{n+1}$ the discrete approximation of $\eta''(u)\psi$ on $\mathcal{V}_{K|L}$ which appears in (2.36), namely:

$$H_{K|L}^{n+1}=\frac{\eta''(\varphi(U_{K,L}^{n+1}))\Psi_K^n+\eta''(\varphi(U_{L,K}^{n+1}))\Psi_L^n}{2}$$

One defines the function $h_{\mathcal{D}}^{\diamond}$ for a.e. $(x,t)\in\Omega\times(0,T)$ by

$$\begin{aligned} h_{\mathcal{D}}^{\diamond}(x,t) &= H_{K|L}^{n+1}, \qquad x\in\mathcal{V}_{K|L},\ t\in(t^n,t^{n+1})\\ h_{\mathcal{D}}^{\diamond}(x,t) &= 0, \qquad x\in\mathcal{V}_{\sigma},\ t\in(t^n,t^{n+1}) \text{ if } \sigma\in\mathcal{E}_{ext}. \end{aligned}$$

Let $\psi_{\mathcal{D}}$ be defined almost everywhere on $\Omega\times(0,T)$ by $\psi_{\mathcal{D}}(x,t)=\Psi_K^n$ for all $(x,t)\in K\times(t^n,t^{n+1})$, for all $K\in\mathcal{T}$ and $n\in[\![0,N]\!]$. The function $\eta''(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}$ tends to $\eta''(\varphi(u))\psi$ in $L^p(\Omega\times(0,T)$ for all $p\in[1,+\infty)$ as $m\to\infty$. Therefore one only needs to compare $h_{\mathcal{D}}^{\diamond}$ and $\eta''(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}$. Since size$(\mathcal{T})$ is small enough, one has

$$\|h_{\mathcal{D}}^{\diamond}-\eta''(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}\|^2_{L^2(\Omega\times(0,T))} \quad = \quad \sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K} m(\mathcal{V}_{K,K|L})(H_{K|L}^{n+1}-\eta''(\varphi(U_K^{n+1}))\Psi_K^n)^2.$$

Let $\varepsilon>0$. The function η'' may be approximated by a function $g\in\mathcal{C}^1(\mathbb{R},\mathbb{R})$ such that $|g(s)-\eta''(s)|<\varepsilon$ for all $s\in[\varphi(U_I),\varphi(U_S)]$. Defining $\tilde H_{K|L}^{n+1}$ and $\tilde h_{\mathcal{D}}^{\diamond}$ using g instead of η'' in the definition of $H_{K|L}^{n+1}$ and $h_{\mathcal{D}}^{\diamond}$ respectively, one has $\|h_{\mathcal{D}}^{\diamond}-\tilde h_{\mathcal{D}}^{\diamond}\|^2_{L^2(\Omega\times(0,T))}\le C_\psi\varepsilon$ and $\|g(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}-\eta''(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}\|^2_{L^2(\Omega\times(0,T))}\le C_\psi\varepsilon$ where $C_\psi\ge 0$ only depends on ψ. Thanks to Young's inequality, one gets

$$\begin{aligned} (\tilde H_{K|L}^{n+1}-g(\varphi(U_K^{n+1}))\Psi_K^n)^2 \quad &\le \quad \left(\max_{s\in[\varphi(U_I),\varphi(U_S)]} g(s)\right)^2(\Psi_K^n-\Psi_L^n)^2\\ &+ \quad \frac{3}{2}\|\psi\|^2_{L^\infty(\Omega\times(0,T))}\left(\max_{s\in[\varphi(U_I),\varphi(U_S)]} g'(s)\right)^2(\varphi(U_K^{n+1})-\varphi(U_L^{n+1}))^2. \quad (2.37) \end{aligned}$$

Using (2.37), the regularity of the function ψ and Estimate (2.22), one gets

$$\|\tilde h_{\mathcal{D}}^{\diamond}-g(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}\|^2_{L^2(\Omega\times(0,T))}\le C(g,\psi,\varphi)\text{size}(\mathcal{T}),$$

where $c(g,\psi,\varphi)\ge 0$ depends only on g,ψ and φ. Hence for size$(\mathcal{T})$ small enough, one has

$$\|\tilde h_{\mathcal{D}}^{\diamond}-g(\varphi(u_{\mathcal{D}}))\psi_{\mathcal{D}}\|^2_{L^2(\Omega\times(0,T))}\le C_\psi\varepsilon,$$

which proves that one can take $m \in \mathbb{N}$ large enough such that

$$\|h^{\diamond}_{\mathcal{D}} - \eta''(u_{\mathcal{D}})\psi_{\mathcal{D}}\|_{L^2(\Omega\times(0,T))} \leq 2C_{\psi}\varepsilon.$$

Hence $h^{\diamond}_{\mathcal{D}_m}$ tends to $\eta''(\varphi(u))\psi$ in $L^2(\Omega \times (0,T))$ as $m \to \infty$.

All the hypotheses of Lemma 2.5.2 (which is statde below) being fulfilled, we may write:

$$-\liminf_{m\to\infty} A4_{\mathcal{D}_m} \leq -\int_0^T\int_\Omega (\nabla\varphi(u)(x,t))^2\eta''(\varphi(u)(x,t))\psi(x,t)dxdt.$$

The proof that u verifies (2.32) is therefore complete.

The same steps are completed in a similar way in order to show that u satisfies (2.33), without the difficult problem of the treatment of η''. This also completes the proof of Theorem 2.5.1. $\square$

To complete the proof of Theorem 2.2.1 there only remains to show the uniqueness of an entropy process solution. This is the aim of Section 2.6.

Lemma 2.5.2 which was used in the above proof is a discrete equivalent of the following continuous classical lemma.

Lemma 2.5.1 *Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of functions of $H^1(\Omega)$ which converges weakly to u in $H^1(\Omega)$ and g a nonnegative function essentially bounded from Ω to $\mathbb{R}$. Then*

$$\int_\Omega (\nabla u(x))^2 g(x)dx \leq \liminf_{n\to\infty}\int_\Omega (\nabla u_n(x))^2 g(x)dx.$$

A discrete version of this lemma is now stated:

Lemma 2.5.2 ("Limit inf" lemma)
Under hypothesis (H), let $g \in L^\infty(\Omega)$ with $g \geq 0$ and let $(\mathcal{T}^{(m)}, \mathcal{E}^{(m)}, (x_K)_{K\in\mathcal{T}^{(m)}})_{m\in\mathbb{N}}$ be a sequence of admissible meshes of Ω in the sense of definition 2.2.1, and $(u^{(m)})_{m\in\mathbb{N}}$ a sequence of piecewise constant functions on Ω such that

- size$(\mathcal{T}^{(m)}) \to 0$ *as $m \to \infty$,*
- *for all $m \in \mathbb{N}$, there exists a set of real values $(G^{(m)}_\sigma)_{\sigma\in\mathcal{E}^{(m)}}$ such that, the functions $g^{(m)}$ defined a.e. on Ω by $g^{(m)}(x) = G^{(m)}_\sigma$ for all $x \in \mathcal{V}_\sigma$ and all $\sigma \in \mathcal{E}_m$, then $g^{(m)} \to g$ in $L^2(\Omega)$ as $m \to \infty$ and the sequence $(g^{(m)})_{m\in\mathbb{N}}$ is bounded in $L^\infty(\Omega)$,*
- *for all $m \in \mathbb{N}$, $u^{(m)}$ is constant in each control volume of $\mathcal{T}^{(m)}$; one denotes by $\mathcal{N}_{\mathcal{T}^{(m)}}$ the value defined by $\mathcal{N}^2_{\mathcal{T}^{(m)}} = \sum_{K|L\in\mathcal{E}^{(m)}_{int}} \tau_{K|L}(u^{(m)}_K - u^{(m)}_L)^2$, where $u^{(m)}_K$ is the constant value of $u^{(m)}$ on $K \in \mathcal{T}^{(m)}$ and $\mathcal{E}^{(m)}_{int}$ is the set of the internal edges of $\mathcal{E}^{(m)}$, and one denotes*

$$D_{\mathcal{T}^{(m)}} = \sum_{K\in\mathcal{T}^{(m)}} \frac{1}{2}\sum_{L\in\mathcal{N}_K} \tau_{K|L} G^{(m)}_{K|L}(u^{(m)}_K - u^{(m)}_L)^2.$$

- *the sequence* $(\mathcal{N}_{\mathcal{T}^{(m)}})_{m\in\mathbb{N}}$ *is bounded,*
- *there exists* $u \in H^1(\Omega)$ *such that* $u^{(m)} \to u$ *in* $L^2(\Omega)$ *as* $m \to \infty$*,*

Then

$$\int_\Omega (\nabla u(x))^2 g(x)dx \quad \leq \quad \liminf_{m\to\infty} D^{(m)} \tag{2.38}$$

Proof. The proof of this lemma is given in [GHM99] in the special case $g = 1$. Let $w \in \mathcal{C}^\infty(\bar{\Omega}, \mathbb{R})$ (the function w is meant to tend to u in $H^1(\Omega)$). Let $m \in \mathbb{N}$, for simplicity of notations, let us write $\mathcal{T} = \mathcal{T}^{(m)}$; let W be the family of values defined by $W_K^{(m)} = w^{(m)}(x_K)$ for $K \in \mathcal{T}$. One compares $Q(g)$ and $Q_\mathcal{T}(g_\mathcal{T}^\diamond)$ defined by

$$\begin{aligned} Q(g) &= \int_\Omega g(x)\nabla u(x)\nabla w(x)dx, \\ Q_\mathcal{T}(g_\mathcal{T}^\diamond) &= \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{L\in\mathcal{N}_K} \tau_{K|L}(u_K - u_L)(W_k - W_L)G_{K|L}. \end{aligned}$$

Since ∇w is bounded, one gets

$$\begin{aligned} |Q(g)| &\leq \|u\|_{H^1(\Omega)}\|\nabla w\|_{L^\infty(\Omega)}\|g\|_{L^2(\Omega)}, \\ |Q_\mathcal{T}(g_\mathcal{T}^\diamond)| &\leq \mathcal{N}_\mathcal{T}\|\nabla w\|_{L^\infty(\Omega)}\|g_\mathcal{T}^\diamond\|_{L^2(\Omega)}. \end{aligned}$$

Let us then prove that:

$$\lim_{m\to\infty} Q_{\mathcal{T}_m}(g_{\mathcal{T}_m}^\diamond) = Q(g). \tag{2.39}$$

Using the density of $\mathcal{C}_c^\infty(\Omega, \mathbb{R})$ in $L^2(\Omega)$, one may assume that $g \in \mathcal{C}_c^\infty(\Omega)$ and that $g_\mathcal{T}^\diamond$ is its natural approximation by the mean values on diamonds for example. For such a regular g, since $u_\mathcal{T}$ converges to u in $L^2(\Omega)$, there exists $\varepsilon(\text{size}(\mathcal{T})$ satisfying $\varepsilon(\text{size}(\mathcal{T})) \longrightarrow 0$ as $\text{size}(\mathcal{T}) \longrightarrow 0$ such that:

$$\begin{aligned} \int_\Omega g(x)\nabla u(x)\nabla w(x)dx &= -\int_\Omega u(x)\,\text{div}(g\nabla w)(x)dx \\ &= -\int_\Omega u_\mathcal{T}(x)\,\text{div}(g\nabla w)(x)dx + \varepsilon(\text{size}(\mathcal{T})). \end{aligned} \tag{2.40}$$

Using the fact that $u_\mathcal{T}$ is piecewise constant, one gets

$$\begin{aligned} -\int_\Omega u_\mathcal{T}(x)\,\text{div}(g\nabla w)(x)dx &= -\sum_{K\in\mathcal{T}} u_K \sum_{L\in\mathcal{N}_K} \int_{K|L} g(x)\nabla w(x)\cdot \mathbf{n}_{K,L}d\gamma(x) \\ &= \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{L\in\mathcal{N}_K} (u_L - u_K) \int_{K|L} g(x)\nabla w(x)\cdot \mathbf{n}_{K,L}d\gamma(x) \\ &= \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{L\in\mathcal{N}_K} (u_L - u_K)\tau_{K|L}G_{K|L}(W_L - W_K) + C_{g,w,\Omega}\mathcal{N}_\mathcal{T}\text{size}(\mathcal{T}). \end{aligned} \tag{2.41}$$

where $C_{g,w,\Omega} \in \mathbb{R}_+$ depends only on g, w and Ω. From the regularity of w and g, one gets (2.39) from (2.40) and (2.41).
By the same proof, replacing u by w, one also has

$$\lim_{\text{size}(\mathcal{T})\to 0} \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{L\in\mathcal{N}_K} \tau_{K|L}(W_L - W_K)^2 G_{K|L} \quad = \quad \int_\Omega (\nabla w(x))^2 g(x) dx.$$

Thanks to the Cauchy-Schwarz inequality, we may write

$$(Q_{\mathcal{T}}(g_{\mathcal{T}}^{\diamond}))^2 \leq D_{\mathcal{T}} \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{L\in\mathcal{N}_K} \tau_{K|L}(W_L - W_K)^2 G_{K|L}. \tag{2.42}$$

Passing to the limit in (2.42) when size($\mathcal{T}$) tends to zero yields

$$(\int_\Omega g(x)\nabla u(x)\nabla w(x)dx)^2 \leq \int_\Omega (\nabla w(x))^2 g(x) dx \quad \liminf_{\text{size}(\mathcal{T})\to 0} D_{\mathcal{T}}. \tag{2.43}$$

Since $C^\infty(\bar{\Omega})$ is dense in $H^1(\Omega)$, one can let $w \to u$ in (2.43), which gives (2.38). $\square$

2.6 Uniqueness of the entropy process solution.

One proves in this section the following theorem.

Theorem 2.6.1 (Uniqueness of the entropy process solution) *Under hypothesis (H), let u and v two entropy process solutions to Problem (2.2)-(2.4) in the sense of definition 2.5.1. Then there exists a unique function $w \in L^\infty(\Omega\times(0,T))$ such that $u(x,t,\alpha) = v(x,t,\beta) = w(x,t)$, for almost every $(x,t,\alpha,\beta) \in \Omega\times(0,T)\times(0,1)\times(0,1)$.*

Proof.
This proof uses on one hand Carrillo's handling of Krushkov entropies, on the other hand the concept of entropy process solution, which allows the use of the theorem of continuity in means, necessary to pass to the limit on mollifiers. Note that the hypothesis (2.5) makes it easier to handle the boundary conditions.

In order to prove Theorem 2.6.1, one defines for all $\varepsilon > 0$ a regularization $S_\varepsilon \in C^1(\mathbb{R},\mathbb{R})$ of the function sign given by

$$\begin{array}{ll} S_\varepsilon(a) = -1, & \forall a \in (-\infty, -\varepsilon], \\ S_\varepsilon(a) = \frac{3\varepsilon^2 a - a^3}{2\ \varepsilon^3}, & \forall a \in [-\varepsilon, \varepsilon], \\ S_\varepsilon(a) = 1, & \forall a \in [\varepsilon, +\infty). \end{array}$$

One defines $\mathbb{R}_\varphi = \{a \in \mathbb{R}, \forall b \in \mathbb{R}\setminus\{a\},\ \varphi(b) \neq \varphi(a)\}$. Note that $\varphi(\mathbb{R}\setminus\mathbb{R}_\varphi)$ is countable, because for all $s \in \varphi(\mathbb{R}\setminus\mathbb{R}_\varphi)$, there exists $(a,b) \in \mathbb{R}^2$ with $a < b$ and $\varphi((a,b)) = \{s\}$, and therefore there exists at least one $r \in \mathbb{Q}$ with $r \in (a,b)$ verifying $\varphi(r) = s$.

Let $\kappa \in \mathbb{R}_\varphi$. Let $\varepsilon > 0$ and let u an entropy process solution. One introduces in (2.32) the function $\eta_{\varepsilon,\kappa}(a) = \int_{\varphi(\kappa)}^a S_\varepsilon(s - \varphi(\kappa))ds$. One defines $\mu_{\varepsilon,\kappa}(a) = \int_\kappa^a \eta'_{\varepsilon,\kappa}(\varphi(s))ds$ and $\nu_{\varepsilon,\kappa}(a) = \int_\kappa^a \eta'_{\varepsilon,\kappa}(\varphi(s))f'(s)ds$, for all $a \in \mathbb{R}$. Using the dominated convergence theorem, one gets for all $a \in \mathbb{R}$ that $\lim_{\varepsilon\longrightarrow 0} \eta_{\varepsilon,\kappa}(a) = |a - \varphi(\kappa)|$, and, since $\kappa \in \mathbb{R}_\varphi$, $\lim\limits_{\varepsilon \longrightarrow 0} \mu_{\varepsilon,\kappa}(a) = |a-\kappa|$ and $\lim\limits_{\varepsilon \longrightarrow 0} \nu_{\varepsilon,\kappa}(a) = f(a\top\kappa) - f(a\perp\kappa)$. One gets for all $\psi \in \mathcal{D}^+(\Omega \times [0,T))$,

$$
\begin{aligned}
&\int_{\Omega\times(0,T)} \begin{bmatrix} \int_0^1 |u(x,t,\alpha)-\kappa| d\alpha\ \psi_t(x,t) \\ + \int_0^1 (f(u(x,t,\alpha)\top\kappa) - f(u(x,t,\alpha)\perp\kappa)) d\alpha\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t) \\ -S_\varepsilon(\varphi(u)(x,t)-\varphi(\kappa))\nabla\varphi(u)(x,t)\cdot\nabla\psi(x,t) \end{bmatrix} dxdt \\
&-\int_{\Omega\times(0,T)} \left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(\kappa))(\nabla\varphi(u))^2(x,t)\psi(x,t)\right] dxdt \\
&+\int_\Omega |u_0(x)-\kappa|\psi(x,0)dx \qquad\qquad \geq A(\varepsilon,u,\kappa,\psi),
\end{aligned}
\tag{2.44}
$$

where for any entropy process solution u, any $\psi \in \mathcal{D}^+(\Omega \times [0,T))$, any $\kappa \in \mathbb{R}_\varphi$ and any $\varepsilon > 0$, $A(\varepsilon,u,\kappa,\psi)$ is defined by

$$
\begin{aligned}
A(\varepsilon,u,\kappa,\psi) &= \int_{\Omega\times(0,T)} \begin{bmatrix} \int_0^1 \Big(|u(x,t,\alpha)-\kappa| - \mu_{\varepsilon,\kappa}(u(x,t,\alpha))\Big) d\alpha\ \psi_t(x,t) + \\ \int_0^1 \Big((f(u(x,t,\alpha)\top\kappa) - f(u(x,t,\alpha)\perp\kappa)) - \nu_{\varepsilon,\kappa}(u(x,t,\alpha))\Big) d\alpha \\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t) \end{bmatrix} dxdt \\
&\quad + \int_\Omega \Big(|u_0(x)-\kappa| - \mu_{\varepsilon,\kappa}(u_0(x))\Big)\psi(x,0)dx.
\end{aligned}
$$

Thanks to the dominated convergence theorem, one has

$$
\lim_{\varepsilon \longrightarrow 0} A(\varepsilon,u,\kappa,\psi) = 0.
$$

This convergence is not uniform w.r.t. κ (even if κ remains bounded), but $A(\varepsilon,u,\kappa,\psi)$ remains bounded (for a given u) if κ ψ, ψ_t and $\nabla\psi$ remain bounded and if the support of ψ remains in a fixed compact set of $\mathbb{R}^d \times [0,T)$.
Using (2.33), one now remarks that, for all $\kappa \in \mathbb{R}$, one has for all $\psi \in \mathcal{D}^+(\Omega \times [0,T))$,

$$
\begin{aligned}
&\int_{\Omega\times(0,T)} \begin{bmatrix} \int_0^1 |u(x,t,\alpha)-\kappa| d\alpha\ \psi_t(x,t) + \\ \int_0^1 (f(u(x,t,\alpha)\top\kappa) - f(u(x,t,\alpha)\perp\kappa)) d\alpha \\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t) \\ -S_\varepsilon(\varphi(u)(x,t)-\varphi(\kappa))\nabla\varphi(u)(x,t)\cdot\nabla\psi(x,t) \end{bmatrix} dxdt \\
&+\int_\Omega |u_0(x)-\kappa|\psi(x,0)dx \qquad\qquad \geq B(\varepsilon,u,\kappa,\psi),
\end{aligned}
\tag{2.45}
$$

where for an entropy process solution u, all $\psi \in \mathcal{D}^+(\Omega \times [0,T))$, all $\kappa \in \mathbb{R}$ and all $\varepsilon > 0$, $B(\varepsilon,u,\kappa,\psi)$ is defined by

$$
B(\varepsilon,u,\kappa,\psi) = \int_{\Omega\times(0,T)} \left[\nabla\Big(|\varphi(u)(x,t)-\varphi(\kappa)| - \eta_{\varepsilon,\kappa}(\varphi(u)(x,t))\Big)\cdot\nabla\psi(x,t)\right] dxdt.
$$

For all $\psi \in \mathcal{D}^+(\Omega \times [0,T))$, one has

$$
B(\varepsilon,u,\kappa,\psi) = -\int_{\Omega\times(0,T)} \left[\Big(|\varphi(u)(x,t)-\varphi(\kappa)| - \eta_{\varepsilon,\kappa}(\varphi(u)(x,t))\Big)\Delta\psi(x,t)\right] dxdt,
$$

and

$$\lim_{\varepsilon \longrightarrow 0} B(\varepsilon, u, \kappa, \psi) = 0,$$

for all $\psi \in \mathcal{D}^+(\Omega \times [0,T))$, $\varepsilon > 0$ and $\kappa \in \mathbb{R}$.

As for the study of A, the quantity $B(\varepsilon, u, \kappa, \psi)$ remains bounded (for a given u) if κ and $\Delta\psi$ remain bounded and if the support of ψ remains in a fixed compact set of $\mathbb{R}^d \times [0,T)$.

Let u and v be two entropy process solutions in the sense of Definition 2.5.1. One defines the sets $E_u = \{(x,t) \in \Omega \times (0,T),\ u(x,t,\alpha) \in \mathbb{R}_\varphi$, for a.e. $\alpha \in (0,1)\}$ and $E_v = \{(x,t) \in \Omega \times (0,T),\ v(x,t,\alpha) \in \mathbb{R}_\varphi$, for a.e. $\alpha \in (0,1)\}$. Indeed, recall that $\varphi(u)$ and $\varphi(v)$ do not depend of $\alpha \in (0,1)$. Then, $\Omega \times (0,T) \setminus E_u = \cup_{s \in \varphi(\mathbb{R}\setminus\mathbb{R}_\varphi)} E_{s,u}$ with $E_{s,u} = \{(x,t) \in \Omega \times (0,T), \varphi(u)(x,t) = s\}$ (the same property is available for v). Let $\xi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R})$ such that, for all $(x,t) \in \Omega \times [0,T)$, $\xi(x,t,\cdot,\cdot) \in \mathcal{D}^+(\Omega \times [0,T))$ and for all $(y,s) \in \Omega \times [0,T)$, $\xi(\cdot,\cdot,y,s) \in \mathcal{D}^+(\Omega \times [0,T))$. One introduces in (2.44), for $(y,s) \in E_v$, and a.e. $\beta \in (0,1)$, $\kappa = v(y,s,\beta)$ and $\psi = \xi(\cdot,\cdot,y,s)$. One integrates the result on $E_v \times (0,1)$. One then gets

$$\begin{aligned}
&\int_{E_v}\int_{\Omega\times(0,T)} \left[\begin{array}{l} \int_0^1\int_0^1 |u(x,t,\alpha) - v(y,s,\beta)| d\alpha d\beta\ \xi_t(x,t,y,s) + \\ \int_0^1\int_0^1 (f(u(x,t,\alpha)\top v(y,s,\beta)) - f(u(x,t,\alpha)\bot v(y,s,\beta))) d\alpha d\beta \\ \mathbf{q}(x,t)\cdot\nabla_x\xi(x,t,y,s) \\ -S_\varepsilon(\varphi(u)(x,t) - \varphi(v)(y,s))\nabla\varphi(u)(x,t)\cdot\nabla_x\xi(x,t,y,s) \end{array}\right] dxdtdyds \\
&- \int_{E_v}\int_{\Omega\times(0,T)} \left[S_\varepsilon'(\varphi(u)(x,t) - \varphi(v)(y,s))(\nabla\varphi(u))^2(x,t)\xi(x,t,y,s)\right] dxdtdyds \\
&+ \int_{E_v}\int_\Omega\int_0^1 |u_0(x) - v(y,s,\beta)|\xi(x,0,y,s) d\beta dxdyds \\
&\geq \int_0^1\int_{E_v} A(\varepsilon, u, v(y,s,\beta), \xi(\cdot,\cdot,y,s)) dydsd\beta.
\end{aligned} \tag{2.46}$$

One introduces in (2.45), for $(y,s) \in \Omega \times (0,T) \setminus E_v$, and any $\beta \in (0,1)$, $\kappa = v(y,s,\beta)$ and $\psi = \xi(\cdot,\cdot,y,s)$. One integrates the result on $(\Omega \times (0,T) \setminus E_v) \times (0,1)$. One then gets

$$\begin{aligned}
&\int_{\Omega\times(0,T)\setminus E_v}\int_{\Omega\times(0,T)} \left[\begin{array}{l} \int_0^1\int_0^1 |u(x,t,\alpha) - v(y,s,\beta)| d\alpha d\beta\ \xi_t(x,t,y,s) \\ +\int_0^1\int_0^1 (f(u(x,t,\alpha)\top v(y,s,\beta)) - f(u(x,t,\alpha)\bot v(y,s,\beta))) d\alpha d\beta \\ \mathbf{q}(x,t)\cdot\nabla_x\xi(x,t,y,s) \\ -S_\varepsilon(\varphi(u)(x,t) - \varphi(v)(y,s))\nabla\varphi(u)(x,t)\cdot\nabla_x\xi(x,t,y,s) \end{array}\right] dxdtdyds \\
&+ \int_{\Omega\times(0,T)\setminus E_v}\int_\Omega\int_0^1 |u_0(x) - v(y,s,\beta)|\xi(x,0,y,s) d\beta dxdyds \\
&\geq \int_0^1\int_{\Omega\times(0,T)\setminus E_v} B(\varepsilon, u, v(y,s,\beta), \xi(\cdot,\cdot,y,s)) dydsd\beta.
\end{aligned} \tag{2.47}$$

Adding (2.46) and (2.47) gives

$$
\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[\begin{array}{l}\int_0^1\int_0^1 |u(x,t,\alpha)-v(y,s,\beta)|d\alpha d\beta\ \xi_t(x,t,y,s)\\ +\int_0^1\int_0^1 (f(u(x,t,\alpha)\top v(y,s,\beta))-f(u(x,t,\alpha)\bot v(y,s,\beta)))d\alpha d\beta\\ \mathbf{q}(x,t)\cdot\nabla_x\xi(x,t,y,s)\\ -S_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))\nabla\varphi(u)(x,t)\cdot\nabla_x\xi(x,t,y,s)\end{array}\right]dxdtdyds
$$
$$
-\int_{E_v}\int_{\Omega\times(0,T)}\left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))(\nabla\varphi(u))^2(x,t)\xi(x,t,y,s)\right]dxdtdyds
$$
$$
+\int_{\Omega\times(0,T)}\int_\Omega\int_0^1 |u_0(x)-v(y,s,\beta)|\xi(x,0,y,s)d\beta dxdyds
$$
$$
\geq\int_0^1\int_{E_v} A(\varepsilon,u,v(y,s,\beta),\xi(\cdot,\cdot,y,s))dydsd\beta+\int_0^1\int_{\Omega\times(0,T)\setminus E_v} B(\varepsilon,u,v(y,s,\beta),\xi(\cdot,\cdot,y,s))dydsd\beta
$$

One now exchanges the roles of u and v, and add the resulting equations. It gives

$$
T_1+T_2+T_3(\varepsilon)+T_4(\varepsilon)+T_5(\varepsilon)\geq T_6(\varepsilon), \tag{2.48}
$$

where

$$
\begin{aligned}
T_1 &= \int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[\begin{array}{l}\int_0^1\int_0^1 |u(x,t,\alpha)-v(y,s,\beta)|d\alpha d\beta\ (\xi_t(x,t,y,s)+\xi_s(x,t,y,s))\\ +\int_0^1\int_0^1 (f(u(x,t,\alpha)\top v(y,s,\beta))-f(u(x,t,\alpha)\bot v(y,s,\beta)))d\alpha d\beta\\ \Big(\mathbf{q}(x,t)\cdot\nabla_x\xi(x,t,y,s)+\mathbf{q}(y,s)\cdot\nabla_y\xi(x,t,y,s)\Big)\end{array}\right]dxdtdyds,\\
T_2 &= \int_{\Omega\times(0,T)}\int_\Omega\int_0^1 |u_0(x)-v(y,s,\beta)|\xi(x,0,y,s)d\beta dxdyds\\
&\quad+\int_{\Omega\times(0,T)}\int_\Omega\int_0^1 |u_0(y)-u(x,t,\alpha)|\xi(x,t,y,0)d\alpha dydxdt,\\
T_3(\varepsilon) &= -\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[\begin{array}{l}S_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))\nabla\varphi(u)(x,t)\cdot\\ (\nabla_x\xi(x,t,y,s)+\nabla_y\xi(x,t,y,s))\end{array}\right]dxdtdyds\\
&\quad-\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[\begin{array}{l}S_\varepsilon(\varphi(v)(y,s)-\varphi(u)(x,t))\nabla\varphi(v)(y,s)\cdot\\ (\nabla_x\xi(x,t,y,s)+\nabla_y\xi(x,t,y,s))\end{array}\right]dxdtdyds,
\end{aligned}
$$

$$
\begin{aligned}
T_4(\varepsilon) &= \int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[S_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))\nabla\varphi(u)(x,t)\cdot\nabla_y\xi(x,t,y,s)\right]dxdtdyds\\
&+\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[S_\varepsilon(\varphi(v)(y,s)-\varphi(u)(x,t))\nabla\varphi(v)(y,s)\cdot\nabla_x\xi(x,t,y,s)\right]dxdtdyds,
\end{aligned} \tag{2.49}
$$

$$
\begin{aligned}
T_5(\varepsilon) &= -\int_{E_v}\int_{\Omega\times(0,T)}\left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))(\nabla\varphi(u))^2(x,t)\xi(x,t,y,s)\right]dxdtdyds\\
&\quad-\int_{\Omega\times(0,T)}\int_{E_u}\left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))(\nabla\varphi(v))^2(y,s)\xi(x,t,y,s)\right]dxdtdyds,
\end{aligned}
$$

and

$$
\begin{aligned}
T_6(\varepsilon) &= \int_0^1 \int_{E_v} A(\varepsilon, u, v(y,s,\beta), \xi(\cdot,\cdot,y,s)) dy ds d\beta \\
&+ \int_0^1 \int_{\Omega\times(0,T)\setminus E_v} B(\varepsilon, u, v(y,s,\beta), \xi(\cdot,\cdot,y,s)) dy ds d\beta \\
&+ \int_0^1 \int_{E_u} A(\varepsilon, v, u(x,t,\alpha), \xi(x,t,\cdot,\cdot)) dx dt d\alpha \\
&+ \int_0^1 \int_{\Omega\times(0,T)\setminus E_u} B(\varepsilon, v, u(x,t,\alpha), \xi(x,t,\cdot,\cdot)) dx dt d\alpha.
\end{aligned}
$$

By an integration by parts in (2.49) and using the fact that ξ vanishes on $\partial\Omega\times(0,T)\times\Omega\times(0,T)$ and on $\Omega\times(0,T)\times\partial\Omega\times(0,T)$ one gets

$$
\begin{aligned}
T_4(\varepsilon) &= \int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)} \left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))\xi(x,t,y,s)\nabla\varphi(u)(x,t)\cdot\nabla\varphi(v)(y,s)\right] dxdtdyds \\
&+\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)} \left[S'_\varepsilon(\varphi(v)(y,s)-\varphi(u)(x,t))\xi(x,t,y,s)\nabla\varphi(v)(y,s)\cdot\nabla\varphi(u)(x,t)\right] dxdtdyds.
\end{aligned}
$$

Recall that $E_{s,u} = \{(x,t)\in\Omega\times(0,T), \varphi(u)(x,t) = s\}$ for all $s\in\mathbb{R}$. One has $\nabla\varphi(u) = 0$ a.e. on $E_{s,u}$ (see [Bré83] for instance). Since $\Omega\times(0,T)\setminus E_u = \cup_{s\in\varphi(\mathbb{R}\setminus\mathbb{R}_\varphi)} E_{s,u}$, and since $\varphi(\mathbb{R}\setminus\mathbb{R}_\varphi)$ is countable, the following equations hold.

$$\nabla\varphi(u) = 0, \text{ a.e. on } \Omega\times(0,T)\setminus E_u$$

and

$$\nabla\varphi(v) = 0, \text{ a.e. on } \Omega\times(0,T)\setminus E_v.$$

It leads to

$$
\begin{aligned}
T_4(\varepsilon) &= \int_{E_u\times E_v} \left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))\xi(x,t,y,s)\nabla\varphi(u)(x,t)\cdot\nabla\varphi(v)(y,s)\right] dxdtdyds \\
&+\int_{E_u\times E_v} \left[S'_\varepsilon(\varphi(v)(y,s)-\varphi(u)(x,t))\xi(x,t,y,s)\nabla\varphi(v)(y,s)\cdot\nabla\varphi(u)(x,t)\right] dxdtdyds
\end{aligned}
$$

and

$$
\begin{aligned}
T_5(\varepsilon) &= -\int_{E_u\times E_v} \left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))(\nabla\varphi(u))^2(x,t)\xi(x,t,y,s)\right] dxdtdyds \\
&-\int_{E_u\times E_v} \left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))(\nabla\varphi(v))^2(y,s)\xi(x,t,y,s)\right] dxdtdyds.
\end{aligned}
$$

Therefore $\forall\varepsilon>0$,

$$
\begin{aligned}
T_4(\varepsilon)+T_5(\varepsilon) &= -\int_{E_v}\int_{E_u} \left[S'_\varepsilon(\varphi(u)(x,t)-\varphi(v)(y,s))\xi(x,t,y,s)\Big(\nabla\varphi(u)(x,t)-\nabla\varphi(v)(y,s)\Big)^2\right] dxdtdyds \\
&\leq 0.
\end{aligned}
$$

One thus gets $\forall\varepsilon>0$,

$$T_1+T_2+T_3(\varepsilon)\geq T_6(\varepsilon). \tag{2.50}$$

One can now let $\varepsilon \longrightarrow 0$ in (2.50). This gives, since $T_6(\varepsilon) \longrightarrow 0$ (thanks to the dominated convergence theorem),

$$\begin{array}{ll}\displaystyle\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[\begin{array}{l}\int_0^1\int_0^1|u(x,t,\alpha)-v(y,s,\beta)|d\alpha d\beta\ (\xi_t(x,t,y,s)+\xi_s(x,t,y,s))+\\ \int_0^1\int_0^1(f(u(x,t,\alpha)\top v(y,s,\beta))-f(u(x,t,\alpha)\bot v(y,s,\beta)))d\alpha d\beta\\ \Big(\mathbf{q}(x,t)\cdot\nabla_x\xi(x,t,y,s)+\mathbf{q}(y,s)\cdot\nabla_y\xi(x,t,y,s)\Big)\\ -(\nabla_x|\varphi(u)(x,t)-\varphi(v)(y,s)|+\nabla_y|\varphi(u)(x,t)-\varphi(v)(y,s)|)\\ \cdot(\nabla_x\xi(x,t,y,s)+\nabla_y\xi(x,t,y,s))\end{array}\right]dxdtdyds & \\ \displaystyle+\int_{\Omega\times(0,T)}\int_\Omega\int_0^1|u_0(x)-v(y,s,\beta)|\xi(x,0,y,s)d\beta dxdyds & \\ \displaystyle+\int_{\Omega\times(0,T)}\int_\Omega\int_0^1|u_0(y)-u(x,t,\alpha)|\xi(x,t,y,0)d\alpha dydxdt & \geq 0.\end{array}\tag{2.51}$$

Now, let us consider the analog of (2.33) for v instead of u, with $\kappa = u_0(x)$ and $\psi(y,s) = \int_s^T \xi(x,0,y,\tau)d\tau$ and integrate the result on $x \in \Omega$. One then gets

$$\begin{array}{ll}\displaystyle\int_\Omega\int_{\Omega\times(0,T)}\left[\begin{array}{l}-\int_0^1|v(y,s,\beta)-u_0(x)|d\beta\ \xi(x,0,y,s)+\\ \int_0^1(f(v(y,s,\beta)\top u_0(x))-f(v(y,s,\beta)\bot u_0(x)))d\beta\ \mathbf{q}(y,s)\cdot\\ \nabla_y\int_s^T\xi(x,0,y,\tau)d\tau\\ -\nabla_y|\varphi(v)(y,s)-\varphi(u_0(x))|\cdot\\ \int_s^T\nabla_y\xi(x,0,y,\tau)d\tau\end{array}\right]dydsdx\quad + & \\ \displaystyle\int_\Omega\int_\Omega|u_0(x)-u_0(y)|\int_0^T\xi(x,0,y,\tau)d\tau dxdy & \geq 0.\end{array}\tag{2.52}$$

A sequence of mollifiers in $\mathbb{R}$ and $\mathbb{R}^d$ is now introduced. Let $\rho \in C_c^\infty(\mathbb{R}^d,\mathbb{R}_+)$ and $\bar\rho \in C_c^\infty(\mathbb{R},\mathbb{R}_+)$ be such that

$$\{x\in\mathbb{R}^d;\ \rho(x)\neq 0\}\subset\{x\in\mathbb{R}^d;\ |x|\leq 1\},$$

$$\{x\in\mathbb{R};\ \bar\rho(x)\neq 0\}\subset[-1,0]$$

and

$$\int_{\mathbb{R}^d}\rho(x)dx=1,\ \int_{\mathbb{R}}\bar\rho(x)dx=1.\tag{2.53}$$

For $n \in \mathbb{N}^\star$, define $\rho_n = n^d\rho(nx)$ for all $x \in \mathbb{R}^d$ and $\bar\rho_n = n\bar\rho(nx)$ for all $x \in \mathbb{R}$.

One sets $\xi(x,t,y,s) = \psi(x,t)\rho_n(x-y)\bar\rho_m(t-s)$, where $\psi \in C_c^\infty(\Omega\times[0,T),\mathbb{R}_+)$ and n and m are large enough to ensure, for all $(x,t)\in\Omega\times[0,T)$, $\xi(x,t,\cdot,\cdot)\in\mathcal{D}^+(\Omega\times[0,T))$ and for all $(y,s)\in\Omega\times[0,T)$, $\xi(\cdot,\cdot,y,s)\in\mathcal{D}^+(\Omega\times[0,T))$. This choice is not symmetrical in (x,t) and (y,s), which gives an easier way to take the limit as $n\longrightarrow\infty$ and $m\longrightarrow\infty$. One gets, from (2.51),

$$\int_{\Omega\times(0,T)}\int_{\Omega\times(0,T)}\left[\begin{array}{l}\rho_n(x-y)\bar\rho_m(t-s)\\ \int_0^1\int_0^1|u(x,t,\alpha)-v(y,s,\beta)|d\alpha d\beta\ \psi_t(x,t)\\ -\int_0^1\int_0^1\left(\begin{array}{l}f(u(x,t,\alpha)\top v(y,s,\beta))\\ -f(u(x,t,\alpha)\bot v(y,s,\beta))\end{array}\right)d\alpha d\beta\\ (\rho_n(x-y)\bar\rho_m(t-s)\mathbf{q}(x,t)\cdot\nabla\psi(x,t)\\ -\psi(x,t)\bar\rho_m(t-s)(\mathbf{q}(x,t)-\mathbf{q}(y,s))\cdot\nabla\rho_n(x-y))\\ -\rho_n(x-y)\bar\rho_m(t-s)(\nabla_x|\varphi(u)(x,t)-\varphi(v)(y,s)|\\ +\nabla_y|\varphi(u)(x,t)-\varphi(v)(y,s)|)\cdot\nabla\psi(x,t)\end{array}\right]dxdtdyds$$
$$+\int_{\Omega\times(0,T)}\int_\Omega\int_0^1|u_0(x)-v(y,s,\beta)|\psi(x,0)\rho_n(x-y)\bar\rho_m(-s)d\beta dxdyds\qquad\geq 0. \tag{2.54}$$

The second of the two initial terms vanishes because of the asymmetric choice of $\bar\rho_m$. Using the same test function in (2.52), at $t=0$, i.e. $\xi(x,0,y,s)=\psi(x,0)\rho_n(x-y)\bar\rho_m(-s)$ and (2.53), we get

$$\int_\Omega\int_{\Omega\times(0,T)}\left[\begin{array}{l}-\int_0^1|v(y,s,\beta)-u_0(x)|d\beta\ \psi(x,0)\rho_n(x-y)\bar\rho_n(-s)\\ -\int_0^1(f(v(y,s,\beta)\top u_0(x))-f(v(y,s,\beta)\bot u_0(x)))d\beta\ \mathbf{q}(y,s)\cdot\\ \psi(x,0)\nabla\rho_n(x-y)\int_s^T\bar\rho_m(-\tau)d\tau\\ +\nabla_y|\varphi(v)(y,s)-\varphi(u_0(x))|\cdot\\ \psi(x,0)\nabla\rho_n(x-y)\int_s^T\bar\rho_m(-\tau)d\tau\end{array}\right]dydsdx$$
$$+\int_\Omega\int_\Omega|u_0(x)-u_0(y)|\psi(x,0)\rho_n(x-y)dxdy\qquad\geq 0. \tag{2.55}$$

One can now add (2.54) and (2.55) let m tend to ∞ and use the theorem of continuity in means. Since the function $s\longrightarrow\int_s^T\bar\rho_m(-\tau)d\tau$ is bounded and tends to zero as $m\longrightarrow\infty$ for all $s\in(0,T)$, one gets

$$\int_\Omega\int_{\Omega\times(0,T)}\left[\begin{array}{l}\rho_n(y-x)\int_0^1\int_0^1|u(x,t,\alpha)-v(y,t,\beta)|d\alpha d\beta\ \psi_t(x,t)\\ +\int_0^1\int_0^1\left(\begin{array}{l}f(u(x,t,\alpha)\top v(y,t,\beta))\\ -f(u(x,t,\alpha)\bot v(y,t,\beta))\end{array}\right)d\alpha d\beta\\ (\rho_n(y-x)\mathbf{q}(x,t)\cdot\nabla\psi(x,t)+\\ \psi(x,t)(\mathbf{q}(y,t)-\mathbf{q}(x,t))\cdot\nabla\rho_n(y-x))\\ -\rho_n(x-y)(\nabla_x|\varphi(u)(x,t)-\varphi(v)(y,t)|\\ +\nabla_y|\varphi(u)(x,t)-\varphi(v)(y,t)|)\cdot\nabla\psi(x,t)\end{array}\right]dxdtdy$$
$$+\int_\Omega\int_\Omega|u_0(x)-u_0(y)|\psi(x,0)\rho_n(x-y)dxdy\qquad\geq 0. \tag{2.56}$$

Remarking that

$$\int_\Omega\int_{\Omega\times(0,T)}\left[\begin{array}{l}\rho_n(x-y)(\nabla_x|\varphi(u)(x,t)-\varphi(v)(y,t)|\\ +\nabla_y|\varphi(u)(x,t)-\varphi(v)(y,t)|)\cdot\nabla\psi(x,t)\end{array}\right]dxdtdy$$
$$=-\int_\Omega\int_{\Omega\times(0,T)}\left[\ \rho_n(x-y)|\varphi(u)(x,t)-\varphi(v)(y,t)|\Delta\psi(x,t)\ \right]dxdtdy,$$

it is possible to let $n\longrightarrow\infty$ in (2.56). Using $\mathrm{div}\mathbf{q}=0$ and the theorem of continuity in means again, one gets

$$\int_{\Omega\times(0,T)}\left[\begin{array}{l}\int_0^1\int_0^1|u(x,t,\alpha)-v(x,t,\beta)|d\alpha d\beta\ \psi_t(x,t)\\ +\int_0^1\int_0^1(f(u(x,t,\alpha)\top v(x,t,\beta))-f(u(x,t,\alpha)\bot v(x,t,\beta)))d\alpha d\beta\\ \mathbf{q}(x,t)\cdot\nabla\psi(x,t)\\ -\nabla|\varphi(u)(x,t)-\varphi(v)(x,t)|\cdot\nabla\psi(x,t)\end{array}\right]dxdt\ \geq 0. \tag{2.57}$$

One notices that (2.57) holds for any $\psi \in H^1(\Omega \times (0,T))$, with $\psi \geq 0$ and $\psi(.,T) = 0$, using a density argument. Therefore one can now take, in (2.57), for ψ the functions $\psi_\varepsilon(x,t) = (T-t)\min(\frac{d(x,\partial\Omega)}{\varepsilon}, 1)$, for $\varepsilon > 0$.
Assume momentarily that for all $w \in H^1_0(\Omega)$ with $w \geq 0$,

$$\liminf_{\varepsilon \longrightarrow 0} \int_\Omega \nabla w(x) \cdot \nabla \min(\frac{d(x,\partial\Omega)}{\varepsilon}, 1)dx \geq 0 \tag{2.58}$$

(The proof of (2.58) is given below).

The expression $\mathbf{q}(x,t) \cdot \nabla \min(\frac{d(x,\partial\Omega)}{\varepsilon}, 1)$ verifies

$$\lim_{\varepsilon \longrightarrow 0} \mathbf{q}(x,t) \cdot \nabla \min(\frac{d(x,\partial\Omega)}{\varepsilon}, 1) = 0, \text{ for a.e. } (x,t) \in \Omega \times (0,T),$$

and under condition (2.5) (and (H5)) remains bounded independently of ε for a.e. $(x,t) \in \Omega \times (0,T)$. Letting $\varepsilon \longrightarrow 0$, (2.57), with $\psi = \psi_\varepsilon$, gives

$$-\int_{\Omega\times(0,T)} \left[\int_0^1 \int_0^1 |u(x,t,\alpha) - v(x,t,\beta)| d\alpha d\beta\right] dxdt \quad \geq 0,$$

which finally proves that $u = v$ and that u is a classical function of space and time (it does not depend on α).

Proof of (2.58)

Let $\varepsilon > 0$. Let $(\partial\Omega_i)_{i=1,\ldots,N}$ be the faces of Ω, $\mathbf{n}_i$ their normal vector outward to Ω, and for $i = 1, \ldots N$, let Ω_i be the subset of Ω such that, for all $x \in \Omega_i$, $d(x,\partial\Omega_i) < \varepsilon$ and $d(x,\partial\Omega_i) < d(x,\partial\Omega_j)$ for all $j \neq i$. One has

$$\int_{\cup_{i=1}^N \Omega_i} \nabla w(x) \cdot \nabla \min(d(x,\partial\Omega)/\varepsilon, 1)dx = \sum_{i=1}^N \int_{\Omega_i} \frac{\nabla w(x) \cdot \mathbf{n}_i}{\varepsilon} dx.$$

For each Ω_i, let $\tilde{\Omega}_i$ be the largest cylinder generated by $\mathbf{n}_i$ included in Ω_i. One denotes by $\partial\Omega'_i$ the face of $\tilde{\Omega}_i$ parallel to $\partial\Omega_i$. Let Ω_ε be defined by $\Omega_\varepsilon = \Omega \setminus \cup_{i=1}^N \tilde{\Omega}_i$. One has $meas(\Omega_\varepsilon) \leq C(\Omega)\varepsilon^2$ and

$$\int_\Omega \nabla w(x) \cdot \nabla \min(d(x,\partial\Omega)/\varepsilon, 1)dx \geq \sum_{i=1}^N \int_{\partial\Omega'_i} \frac{w(x)}{\varepsilon} d\gamma(x) - \int_{\Omega_\varepsilon} \frac{|\nabla w(x)|}{\varepsilon} dx.$$

Thanks to the Cauchy-Schwarz inequality, one gets

$$(\int_{\Omega_\varepsilon} |\nabla w(x)| dx)^2 \leq \text{meas}(\Omega_\varepsilon) \int_{\Omega_\varepsilon} (\nabla w(x))^2 dx.$$

One concludes, using $\displaystyle\lim_{\varepsilon \longrightarrow 0} \int_{\Omega_\varepsilon} (\nabla w(x))^2 dx = 0$.

Remark 2.6.1 *Inequality (2.58) could also be proved in the case where Ω is regular instead of polygonal, with a slightly different method. Let $\Omega_\varepsilon = \{x \in \Omega, d(x,\partial\Omega) < \varepsilon\}$ and let $\partial\Omega'_\varepsilon$ be the other face of Ω_ε. The normal vector to $\partial\Omega'_\varepsilon$ at any point x is equal to $\nabla d(x,\partial\Omega)$. Therefore one has*

$$\int_\Omega \nabla w(x)\cdot\nabla \min(d(x,\partial\Omega)/\varepsilon,1)dx = \int_{\partial\Omega'_\varepsilon}\frac{w(x)}{\varepsilon}d\gamma(x) - \int_{\Omega_\varepsilon} w(x)\frac{\Delta d(x,\partial\Omega)}{\varepsilon}dx.$$

Since Hardy's inequality leads to

$$\int_{\Omega_\varepsilon}\left(\frac{w(x)}{d(x,\partial\Omega)}\right)^2 dx \le C(\Omega)\int_{\Omega_\varepsilon}(\nabla w(x))^2 dx,$$

one concludes using $m_{\varepsilon\longrightarrow 0}\int_{\Omega_\varepsilon}(\nabla w(x))^2dx = 0.$

□

2.7 Conclusion

Let us finally prove the convergence theorem by way of contradiction:
Assume that the convergence stated in the Theorem 2.2.1 does not hold. Then there exist $\varepsilon > 0$, $p\in[1,+\infty)$ and a sequence $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ such that $\|u_{\mathcal{D}_m}-u\|_{L^p(\Omega\times(0,T)}\ge\varepsilon$, for any $m\in\mathbb{N}$. Then by Theorem 2.5.1, there exists a subsequence of the sequence $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$, still denoted by $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ which converges to an entropy process solution of (2.2)-(2.4) . By Theorem 2.6.1 this entropy process solution is the unique entropy weak solution to (2.2)-(2.4) , and from Lemma 2.7.1 which is stated below, the convergence of $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ is strong in any $L^q(\Omega\times(0,T))$. This is in contradiction with the fact that $\|u_{\mathcal{D}_m}-u\|_{L^p(\Omega\times(0,T)}\ge\varepsilon$, for any $m\in\mathbb{N}$.

Lemma 2.7.1 *Let Q be a Borel subset of $\mathbb{R}^k$ and let $(u_n)_{n\in\mathbb{N}}\subset L^\infty(Q)$ be such that u_n converges to $u\in L^\infty(Q\times(0,1))$ in the nonlinear weak star sense where u does not depend on α, then $(u_n)_{n\in\mathbb{N}}$ converges to u in $L^p_{loc}(Q)$ for any $p\in[1,\infty)$.*

Proof. Let K be a compact subset of Q, since u_n converges to u in the nonlinear weak star sense , one has

$$\int_K |u_n(x)-u(x)|^2dx = \int_K u_n^2(x)dx - 2\int_K u_n(x)u(x)dx + \int_K u(x)^2dx \longrightarrow 0 \text{ as } n\longrightarrow+\infty;$$

since K is bounded, one also has:

$$\int_K |u_n(x)-u(x)|^p dx \longrightarrow 0 \text{ as } n\longrightarrow+\infty,\ \forall p\in[1,2]$$

and since the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in $L^\infty(Q)$,

$$\int_K |u_n(x)-u(x)|^p dx \longrightarrow 0 \text{ as } n\longrightarrow+\infty,\ \forall p>2.$$

□

Remark 2.7.1 *An interesting (and open to our knowledge) question is to find the convergence rate of the finite volume approximations. In the case of a pure hyperbolic equation, i.e. $\varphi=0$, it was proven by several authors (under varying assumptions, see e.g. [CCL95], [Vil94], [EGGH98], [CH99]) that the error between the approximate finite volume solution and the entropy weak solution is of order less than $h^{1/4}$ where h is the size of the mesh, under a usual CFL condition for the explicit schemes which are considered* *in [CCL95], [Vil94], [EGGH98], [CH99], and of order less than* $h^{1/4} + k^{1/2}$ *where* k *is the time step in the case of the implicit scheme considered in [EGGH98]. However, it is also known that these estimates are not sharp, since numerically the order of the error behaves as* $1/2$.

In the case of a pure linear parabolic equation, estimates of order 1 were obtained in [Her96] (see also [EGH00b])

We made a first attempt in the direction of an error estimate in the case of the present degenerate parabolic equation by looking at the analogous continuous problem [EGH00a]: let u_ε *be the unique solution to*

$$u_t(x,t) + \mathrm{div}\Big(\mathbf{q}\ f(u)\Big)(x,t) - \Delta\varphi(u)(x,t) - \varepsilon\Delta u(x,t) = 0,\ \textit{for}\ (x,t) \in \Omega \times (0,T),$$

with initial condition (2.3) and boundary condition (2.4) and let u *be the unique entropy weak solution solution of (2.2)-(2.4) , then under assumptions (H), we are able to prove that* $\|u_\varepsilon - u\|_{L^1(Q_T)} \le C\varepsilon^{1/5}$ *where* $C \in \mathbb{R}_+$ *depends only on the data. This estimate is however probably not optimal and we have not yet been able to transcribe its proof to the discrete setting (the term* $-\varepsilon\Delta u$ *being the continuous diffusive representation of the diffusive perturbation introduced by the finite volume scheme).*

2.8 A numerical example

We finally present some numerical results which we obtained by implementing the scheme which was studied above in a prototype code.

The domain Ω is the unit square $(0,1)\times(0,1)$. We define two subregions $\Omega_1 = (0.1, 0.3)\times(0.4, 0.6)$ and $\Omega_2 = (0.7, 0.9)\times(0.4, 0.6)$. The initial data is given by 0.5 in $\Omega \setminus (\Omega_1 \cup \Omega_2)$, 1 in Ω_1 and 0 in Ω_2. It is represented on upper left corner of the figure below. The boundary value is the constant 0.5.

The function φ is defined by $\varphi(s) = 0$ if $s \in [0, 0.5]$ and $\varphi(s) = 0.2(s - 0.5)$ if $s \in [0.5, 1]$, so that the diffusion effect only takes place in the areas where the saturation u is greater than 0.5 . The function f is defined by $f(s) = s$ and the field $\mathbf{q}$ is defined by $\mathbf{q}(x,y) = (10(x - x^2)(1-2y), -10(y-y^2)(1-2x))$. Hence there is a linear rotating convective transport.

We define a coarse mesh of 14 admissible triangles on the unit square, from which we obtain a fine mesh of roughly 12600 triangles by refining these 14 triangles uniformly 30 times. This fine mesh is used for the computations.

The figure below presents the obtained results at times 0.000, 0.007, 0.028 and 0.112. The black points correspond to the value 1, the white ones to the value 0, with a continuous hot-colors scale of between these values. One observes that the initial value 0 is transported, only modified by the numerical diffusion due to the convective upstream weighting, and that, on the contrary, the initial value 1 is rapidly smoothed, due to the effect of the parabolic term which is active on the range $[0.5, 1]$.

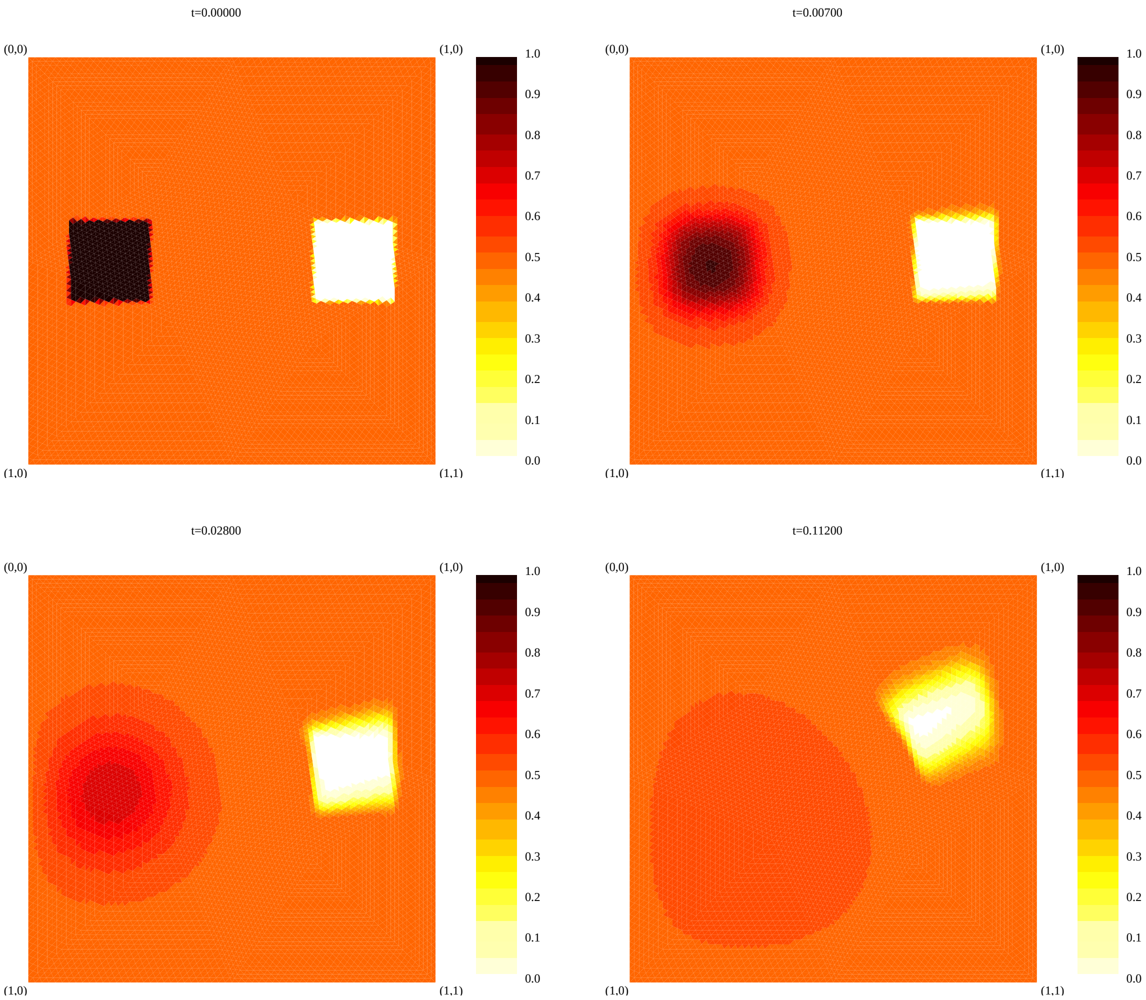


Fig 1.
Computed solution at time $t = 0$ (initial condition), $t = 0.007$, $t = 0.028$ and $t = 0.112$.

Chapitre 3

Equation parabolique hyperbolique dégénérée avec condition de Dirichlet non homogène dans le cas général

Résumé

Cet article (travail commun avec Julien Vovelle) est consacré à l'étude d'une méthode de discrétisation pour une équation parabolique hyperbolique fortement dégénérée sur un domaine borné. La donnée initiale u_0 est une fonction mesurable bornée, la condition au bord $\bar{u}$ est une fonction continue, vérifiant certaines hypothèses de régularité que l'on précise dans l'article. On introduit une notion de solution faible entropique généralisée pour le problème continu (3.1) et on prouve que la solution est unique. On démontre également que l'approximation numérique construite par un schéma volumes finis implicite converge vers une solution faible entropique dans $L^p(Q)$, pour tout $p \in [1, +\infty[$, ce qui fournit en même temps une preuve d'existence.

Le problème considéré constitue une généralisation des cas traités dans [Car99], [Vov00] ainsi que du cas traité au chapitre 2. La grande originalité de ce travail réside dans la formulation faible entropique, qui étend les formulations intégrales précédentes, ainsi que dans la preuve d'unicité. Néanmoins, le traitement des conditions au bord pour la convergence est plus difficile que dans les cas traités aux chapitres 1 et 2 parce que le support en x des fonctions test n'est pas inclus dans le domaine Ω. La formulation intégrale du problème que nous donnons facilite le travail de la preuve de convergence, et donc de la preuve d'existence, par rapport à d'autres formulations basées sur des limites de flux entropiques au bord ([Ott96a],[MPT00]). Des restrictions sur la régularité de la donnée au bord, précisées dans l'hypothèse (H6), sont nécessaires techniquement dans la preuve d'unicité, mais la formulation reste valable pour des données au bord plus générales.

A finite volume method for parabolic degenerate problems with general Dirichlet boundary conditions

Abstract

This paper is devoted to the study of the finite volume methods used in the discretization of conservation laws defined on bounded domains. General assumptions are made on the data: the initial condition and the boundary condition are supposed to be measurable bounded functions. Using a generalized notion of solution to the continuous problem (namely the notion of entropy process solution, see [EGH00b]) and a uniqueness result on this solution, we prove that the numerical solution converges to the entropy weak solution of the continuous problem in L^p_{loc} for every $p \in [1, +\infty)$. This also yields a new proof of the existence of an entropy weak solution.

3.1 Introduction

Let Ω be an open bounded polyhedral subset of $\mathbb{R}^d$ and $T \in \mathbb{R}^*_+$. Let us denote by Q the set $(0,T) \times \Omega$ and by Σ the set $(0,T) \times \partial\Omega$.

We consider the following parabolic-hyperbolic problem:

$$\begin{cases} u_t(t,x) + \operatorname{div}(F(t,x,u)) - \Delta(\varphi(u))(t,x) = 0\,,\ (t,x) \in Q \\ u(0,x) = u_0\,,\ x \in \Omega \\ u(t,x) = \bar{u}(t,x)\,,\ (t,x) \in \Sigma\,. \end{cases} \tag{3.1}$$

When the function φ is strictly increasing, this problem is of parabolic type and it is well known that there exists a unique weak solution. In the case where φ is equal to zero, then 3.1 is a hyperbolic nonlinear problem with non homogeneous Dirichlet boundary conditions, and there is no more uniqueness of a weak solution, besides, the meaning of the boundary condition is not so clear. Nevertheless, by the use of an entropy formulation of the equation and of the boundary conditions, we can also prove a theorem of existence and uniqueness. In this paper we are interested in the more general case where φ is a Lipschitz continuous non decreasing function, which means in particular that it can be constant on intervals with non zero measure. Thus the physical domains where the problem is of parabolic type or of hyperbolic type depend on the values of the solution itself and we must be able to treat both types of phenomena in a unique formulation.

We make the same assumptions on the data as in the hyperbolic case, except for the boundary condition $\bar{u}$ that is supposed to be somewhat more regular.

Hypotheses (H):

$(H1)$ $F : (t,x,s) \mapsto F(t,x,s) \in C^1(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R})$,
$\dfrac{\partial F}{\partial s}$ locally Lipschitz continuous uniformly with respect to (x,t),

$$\operatorname{div}_x F(t,x,s) = \sum_{i=1}^{N} \frac{\partial F_i}{\partial x_i}(t,x,s) = 0, \forall\ (t,x,s) \in \mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R},$$

$(H2)$ $\varphi : s \mapsto \varphi(s)$ is a nondecreasing Lipschitz continuous function,

$(H3)$ $u_0 : x \mapsto u_0(x) \in L^\infty(\Omega)$,

and $(H4)$ $\bar{u} : (x,t) \to \bar{u}(x,t) \in C(\Sigma)$,

$(H5)$ *$\bar{u}$ is the trace of a function $\bar{u} \in L^\infty(Q)$,* such that $\nabla \bar{u} \in L^2(Q), \bar{u}_t \in L^1(Q)$ and

$(H6)$ *for any local extension $\overline{u}_\Sigma$ of $\overline{u}$ $\Delta(\varphi(\overline{u}_\Sigma))$ is in L^1_{loc} (with the notations defined in Subsection 3.4.1).*

Under these assumptions there exists $(A,B) \in \mathbb{R}^2$, such that

$$A \leq \min(ess \inf_{\Omega}(u_0), ess \inf_{Q}(\bar{u})) \leq \max(ess \sup_{\Omega}(u_0), ess \sup_{Q}(\bar{u})) \leq B \tag{3.2}$$

Theses hypotheses allow for example to treat the classical cases $F(x,t,s) = F(s)$ and $F(x,t,s) = \mathbf{q}(x,t)f(s)$ with $\text{div}(\mathbf{q}) = 0$. The reason why we make Assumptions (H4)(H5)(H6) is not clear at a first look nor in the definition of the problem, nor in the entropy formulation (see Definition 3.3.1) itself. Assumption (H5) is used in Section 3.7 to obtain a priori estimates on the discrete solution which is the main ingredient to obtain compactness properties on the sequence of approximate solutions. On the contrary Assumption (H6) which signifies $\Delta\varphi(\bar{u}) \in L^1(Q)$ is only a sufficient assumption to prove Lemma 3.4.1 and we do not know if it can be weakened.

To our knowledge, the problem that we deal with has never been treated before in the literature. The only work that handles the degenerate parabolic problem with non homogeneous boundary conditions is the recent work [MPT00]. Nevertheless some particular cases have been fully treated. In [EGHM00], the authors prove the convergence of finite volumes method in the case where $F(x,t,s) = \mathbf{q}(x,t)f(s)$, $\text{div}(\mathbf{q}) = 0$ with the boundary condition $\mathbf{q} \cdot \mathbf{n} = 0$. This work follows the methods of Carrillo in [Car99] who only deals with homogeneous boundary conditions. In [Vov00], the author proves the convergence of finite volumes method in the case where $\varphi = 0$, adapting the ideas of F.Otto [Ott96a]. In another direction a comparison between two solutions with different diffusive terms using semi-group theory has also been done by Cockburn and Gripenberg [CG99] in the case where $\Omega = \mathbb{R}^d$.

The difficulties involved in this work are of two types. First, we explore the difficult concept of entropy formulations of problems of hyperbolic type with non homogeneous boundary conditions. In that direction, one can remark that the entropy formulation given in Definition 3.3.1 generalizes all the definitions given in the chronology by S.N. Kruzkhov, J.Carrillo, and F. Otto. A large part of the ideas of the proof of uniqueness were already contained in the work of J.Vovelle [Vov00], nevertheless, many ingredients are specific to the case we deal with here. The arguments of the proof are inspired from the article of C.Mascia, A.Porratta and A.Terracina [MPT00]. The originality of our work is the definition of a weak entropy solution using only an integral formulation. This is the integral formulation that allows us to obtain existence of a solution from the convergence of the scheme. As in [Vov00] or [EGHM00], we use the useful tool of measure valued solutions introduced by DiPerna [DiP85] [EGH00b], that we call entropy process solutions, to deal with nonlinear convergence.

In this paper, we first define entropy process solutions and prove the fundamental theorem of uniqueness 3.4.1. Then we deal with the weak convergence of a finite volume scheme to an entropy process solution that gives in the same way the existence of an entropy process solution. Then, by using measure valued functions properties and the uniqueness theorem 3.4.1 we get the strong convergence of the approximate solution obtained by the scheme defined in Section 3.5.2 to an entropy solution of Problem (3.1) that is to say a function defined almost everywhere on $\Omega \times (0,T)$ satisfying Definition 3.2.1.

3.2 Entropy weak solution

Here, as in the study of purely hyperbolic problems, the concept of weak solutions is not sufficient, as uniqueness of such a solution may fail. Thus, we turn to the notion of weak entropy solution. The entropy-flux pairs considered in the definition of this solution are the "semi-Kruzkov entropy-flux pairs" (see [Car99], [Ser96], [Vov00]).

Notations: set

$$s\top\kappa = \begin{cases} s \textit{ if } s > \kappa \\ \kappa \textit{ if } s \leq \kappa \end{cases}, \quad s\bot\kappa = \begin{cases} s \textit{ if } s < \kappa \\ \kappa \textit{ if } s \geq \kappa \end{cases}, \quad s^+ = s\top 0\,, \quad s^- = -(s\bot 0).$$

Then, the semi-Kruzkov entropies are defined by

$$\eta_\kappa^+(s) = (s-\kappa)^+ = s\top\kappa - \kappa \quad \text{and} \quad \eta_\kappa^-(s) = (s-\kappa)^- = \kappa - s\bot\kappa\,, \tag{3.3}$$

while the entropy fluxes associated to these entropies are defined by

$$\Phi_\kappa^+(t,x,s) = F(t,x,s\top\kappa) - F(t,x,\kappa) \quad \text{and} \quad \Phi_\kappa^-(t,x,s) = F(t,x,\kappa) - F(t,x,s\bot\kappa)\,. \tag{3.4}$$

In the case where we consider κ as a variable, for example in the doubling variable technique of Kruzkov, we denote by

$$\Phi^+(x,t,s,\kappa) = \Phi_\kappa^+(t,x,s) \quad \text{and} \quad \Phi^-(x,t,s,\kappa) = \Phi_\kappa^-(t,x,s)$$

The functions Φ^+ and Φ^- are not symmetric in s and κ. Note that the classical entropy flux Φ of Kruzkov associated to the entropy $s \mapsto |s-\kappa|$ which is defined by

$$\varphi(t,x,s,\kappa) = F(t,x,s\top\kappa) - F(t,x,s\bot\kappa)$$

is symmetric with respect to s and k, this not the case for the functions Φ^+ and Φ^-.

Remark 3.2.1 *Note that η_κ^+ and Φ_κ^+ can also be expressed differently using the function* sgn^+*:*

$$\eta_\kappa^+(s) = \mathrm{sgn}^+(u-s) \quad \textit{and} \quad \Phi_\kappa^+(t,x,s) = \mathrm{sgn}^+(u-s)(F(t,x,u) - F(t,x,s)).$$

A similar property for η_κ^- and Φ_κ^- with the function sgn^- *also holds.*

Definition 3.2.1 (Entropy weak solution) *A function u of $L^\infty(Q)$ is said to be an entropy weak solution to problem (3.1) if it is a weak solution of the problem (3.1), that is to say:*

$$\begin{cases} \varphi(u) - \varphi(\overline{u}) \in L^2(0,T;H_0^1(\Omega))\,, \\ \\ \textit{and, for all } \theta \in \mathcal{C}_c^\infty([\,0,T) \times \Omega)\,, \\ \\ \displaystyle\iint_Q \Big[u(t,x)\,\theta_t(t,x) + (F(t,x,u(t,x) - \nabla(\varphi(u)(t,x))) \cdot \nabla\theta(t,x)\Big]\,dx\,dt \,+ \int_\Omega u_0\,\theta(0,x)\,dx \,= 0\,, \end{cases} \tag{3.5}$$

and if u satisfies the following entropy inequalities:

1. *for all $\kappa \in \mathrm{I\!R}$, for all $\psi \in \mathcal{C}_c^\infty([\,0,T) \times \mathrm{I\!R}^d)$ such that $\psi \geq 0$ and $\mathrm{sgn}^+(\overline{u}-\kappa)\psi = 0$ on Σ,*

$$\begin{aligned} &\iint_Q \left(\eta_\kappa^+(u(t,x))\,\psi_t(t,x) + \Phi_\kappa^+(t,x,u(t,x)) \cdot \nabla\psi(t,x)\right)\,dx\,dt \\ &- \iint_Q \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^+ \cdot \nabla\psi(t,x)\,dx\,dt \,+ \int_\Omega \eta_\kappa^+(u_0)\,\varphi(0,x)\,dx \,\geq 0\,, \end{aligned} \tag{3.6}$$

2. *for all $\kappa \in \mathrm{I\!R}$, for all $\psi \in \mathcal{C}_c^\infty([\,0,T) \times \mathrm{I\!R}^d)$ such that $\psi \geq 0$ and $\mathrm{sgn}^-(\overline{u}-\kappa)\psi = 0$ on Σ,*

$$\begin{aligned} &\iint_Q \left(\eta_\kappa^-(u(t,x)\,\psi_t(t,x) + \Phi_\kappa^-(t,x,u(t,x)) \cdot \nabla\psi(t,x)\right)\,dx\,dt \\ &- \iint_Q \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^- \cdot \nabla\psi(t,x)\,dx\,dt \,+ \int_\Omega \eta_\kappa^-(u_0)\,\varphi(0,x)\,dx \,\geq 0\,, \end{aligned} \tag{3.7}$$

Notice that the weak equation exposed in (3.5) is superfluous, for it is a consequence of Equations (3.6) and (3.7). However, if the function φ were (strictly) increasing, Equation (3.5) would be enough to define a notion of solution of Problem (3.1) for which existence and uniqueness hold: in that case, Problem (3.1) would merely be a non-linear parabolic problem.

Notice also that the class of Kruzkov semi entropy-flux pairs is wide enough to ensure the uniqueness of the weak entropy solution. We refer to [Vov00] and [Ott96a][MNRR96] for explanations in the case of the hyperbolic problem. Moreover, the possibility to define numerical entropy fluxes associated to these entropies is one of the keys of the result of convergence of the scheme.

The class of test-functions considered in the previous definition depends on the boundary condition $\overline{u}$. To justify this choice of test functions, one can remark that the condition $\mathrm{sgn}^+(u-\kappa)\psi = 0$ on Σ does not allows to avoid having, in the inequality (3.6), the boundary integrals yielded by the integration by part of the hyperbolic term. To explain why we need exactly this class of test functions one can compare our formulation to the formulation given in [Vov00] in the hyperbolic case. This class has to be wide enough to ensure the uniqueness. In particular, this formulation is only valid if the boundary condition is continuous even in the hyperbolic case and the importance of the continuity of $\bar{u}$ clearly appears in the proof of proposition 3.4.1.

3.3 Entropy process solution

The proof of the existence of a weak entropy solution to Problem (3.1) lies in the study of the numerical solution u_D defined by the finite volume method applied to (3.1) (see Section 3.5.2). Theorem 3.8.1 states that the numerical solution satisfies the following approximate entropy inequalities:

$$\left\{ \begin{array}{l} \forall\kappa \in \mathbb{R}\,,\ \text{for all}\ \psi \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^d)\ \text{such that}\ \psi \geq 0\ \text{and}\ \mathrm{sgn}^+(\overline{u}-\kappa)\psi = 0\ \text{on}\ \Sigma\,, \\ \displaystyle\int_0^T\!\!\int_\Omega (\eta_\kappa^+(u_D)\psi_t + \Phi_\kappa^+(t,x,u_D)\cdot\nabla\psi + \eta_{\varphi(\kappa)}^+(\varphi(u_D))\Delta\psi) - \int_\Sigma \eta_{\varphi(\kappa)}^+(\varphi(\bar{u}))\nabla\psi\cdot\mathbf{n} \\ \hfill \displaystyle + \int_\Omega \eta_\kappa^+(u_0)\psi(0) \geq -\mathcal{E}_D^+(\psi) \end{array} \right. \tag{3.8}$$

where for a given φ, $\mathcal{E}_D^+(\psi) \longrightarrow 0$ when the size of the discretization tends to zero. The same result holds when the entropy-flux pair $(\eta_\kappa^-, \Phi_\kappa^-)$ is considered.

The numerical approximate solution u_D is known to be bounded in $L^\infty(Q)$. An estimate on the discrete H^1-norm of $\varphi(u_D)$ also holds (see Propositions 3.6.1 and 3.7.1). Nevertheless it is not enough to pass to the limit in inequality (3.8). Thus, owing to the non-linearity of the equation and to the lack of estimates on the approximate solution, we have to turn to the notion of measure-valued solutions (see DiPerna, [DiP85], Szepessy, [Sze89]) or, equivalently, to the notion of entropy process solution defined by Eymard, Gallouët and Herbin [EGH00b]. The interest of this notion lies in the following result, which generalizes the notion of weak-$\star$ convergence in L^∞ and frees oneself from the difficulties of non-linearities.

Theorem 3.3.1 (Non-linear weak star convergence) *Let $\mathcal{O}$ be a Borel subset of $\mathbb{R}^m$, let R be positive and (u^n) be a sequence of $L^\infty(\mathcal{O})$ such that, for all $n \in \mathbb{N}$, $||u^n||_{L^\infty} \leq R$. Then there exists a sub-sequence still denoted by (u^n) and $\mu \in L^\infty(\mathcal{O}\times(0,1))$ such that:*

$$\forall g \in \mathcal{C}(\mathbb{R})\,,\ g(u^n) \longrightarrow \int_0^1 g(\mu(.,\alpha))\,d\alpha\ \ \textit{in}\ L^\infty(\mathcal{O})\ \ \textit{weak-}\star\ .$$

Now the notion of entropy process solution can be defined.

Definition 3.3.1 (Weak entropy process solution) *Let u be in $L^\infty(Q \times (0,1))$. The function u is said to be an entropy process solution to problem (3.1) if:*

$$\varphi(u) - \varphi(\overline{u}) \in L^2(0,T; H_0^1(\Omega))\,, \tag{3.9}$$

and if u satisfies the following entropy inequalities:

1. *for all $\kappa \in \mathbb{R}$, for all $\psi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^d)$ such that $\psi \geq 0$ and $\operatorname{sgn}^+(\overline{u} - \kappa)\psi = 0$ on Σ,*

$$\iint_Q \int_0^1 \big(\eta_\kappa^+(u(t,x,\alpha))\,\psi_t(t,x) + \Phi_\kappa^+(t,x,u(t,x,\alpha)) \cdot \nabla\psi(t,x)\big)\, d\alpha\, dx\, dt$$
$$- \iint_Q \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^+ \cdot \nabla\psi(t,x)\, dx\, dt + \int_\Omega \eta_\kappa^+(u_0)\,\varphi(0,x)\, dx \geq 0\,,$$

2. *for all $\kappa \in \mathbb{R}$, for all $\psi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^d)$ such that $\psi \geq 0$ and $\operatorname{sgn}^-(\overline{u} - \kappa)\psi = 0$ on Σ,*

$$\iint_Q \int_0^1 \big(\eta_\kappa^-(u(t,x,\alpha))\,\psi_t(t,x) + \Phi_\kappa^-(t,x,u(t,x,\alpha)) \cdot \nabla\psi(t,x)\big)\, d\alpha\, dx\, dt$$
$$- \iint_Q \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^- \cdot \nabla\psi(t,x)\, dx\, dt + \int_\Omega \eta_\kappa^-(u_0)\,\varphi(0,x)\, dx \geq 0\,,$$

Notice that if u is an entropy process solution of Problem (3.1) it satisfies Condition (3.9), in particular, the function $\varphi(u)$ does not depend on the last variable α. It is denoted by $\varphi(u)(t,x)$.
Notation: in order to lighten the writing, the integration of a function h on the domain $Q \times (0,1)$ will be denoted by

$$\int_{Q^1} h(t,x,\alpha)\, d\alpha\, dx\, dt = \iint_Q \int_0^1 h(t,x,\alpha)\, d\alpha\, dx\, dt\,.$$

Lemma 3.3.1 (Non degenerate zone entropy inequalities) *Let $u \in L^\infty(Q \times (0,1))$ be a weak entropy process solution to Problem (3.1). Then the function u satisfies the following entropy inequalities: for all convex functions $\mu \in \mathcal{C}^2(\mathbb{R})$, for all $\psi \in \mathcal{C}_c^\infty([0,T) \times \Omega)$ such that $\psi \geq 0$,*

$$\int_{Q^1} \eta(u(t,x,\alpha))\,\psi_t(t,x) + [\nu(t,x,u(t,x,\alpha)) - \nabla\mu(\varphi(u)(t,x))] \cdot \nabla\psi(t,x)\, d\alpha\, dx\, dt$$
$$- \iint_Q \mu''(\varphi(u))\,(\nabla\varphi(u))^2\,\psi(t,x)\, dx\, dt + \int_\Omega \eta(u_0)\,\psi(0,x)\, dx \geq 0\,,$$

where $\eta' = \mu'(\varphi)$ and $\dfrac{\partial\nu}{\partial s} = \mu'(\varphi)\dfrac{\partial f}{\partial s}$

Lemma 3.3.1 does not concern the boundary of Ω since the test functions have compact support in Ω and the conclusion is also true if u is only a weak process solution. See the proof in ([Car99], Lemma 4.8).

3.4 Uniqueness of the entropy process solution

Theorem 3.4.1 (Uniqueness of the entropy process solution) *Let u, $v \in L^\infty(Q \times (0,1))$ be two entropy process solutions of Problem 3.1 in the sense of Definition 3.3.1 . Then there exists $w \in L^\infty(Q)$ such that:*

$$u(t,x,\alpha) = w(t,x) = v(t,x,\beta) \text{ for a.e. } (t,x,\alpha,\beta) \in Q \times (0,1)^2 .$$

Corollary 3.4.1 (Uniqueness of the weak entropy solution) *The problem (3.1) admits at most one weak entropy solution.*

3.4.1 Proof of Theorem 3.4.1, definitions and notations

We suppose Ω to be a polyhedral open subset of $\mathbb{R}^d$. Note that the following proof is still correct if Ω is any strong Lipschitz open subset of $\mathbb{R}^d$. In that case, there exists a finite open cover $(B_\nu)_{0,N}$ of $\overline{\Omega}$ and a partition of unity $(\lambda_\nu)_{0,N}$ on $\overline{\Omega}$ subordinate to $(B_\nu)_{0,N}$ such that, up to a change of coordinates represented by an orthogonal matrix A_ν, the set $\Omega \cap B_\nu$ is the epigraph of a Lipschitz continuous function $f_\nu : \mathbb{R}^{d-1} \mapsto \mathbb{R}$, that is to say:

$$\Omega \cap B_\nu = \{x \in B_\nu \; ; \; (A_\nu\, x)_d > f_\nu(\overline{A_\nu\, x})\}$$

and

$$\partial\Omega \cap B_\nu = \{x \in B_\nu \; ; \; (A_\nu\, x)_d = f_\nu(\overline{A_\nu\, x})\} ,$$

where $\overline{y}$ stands for $(y_i)_{1,d-1}$ if $y \in \mathbb{R}^d$.

Until the end of the proof of Theorem 3.4.1, the problem will be localized with the help of a function λ_ν. We drop the index ν and, for the sake of clarity, suppose that the change of coordinates is trivial: $A = I_d$.

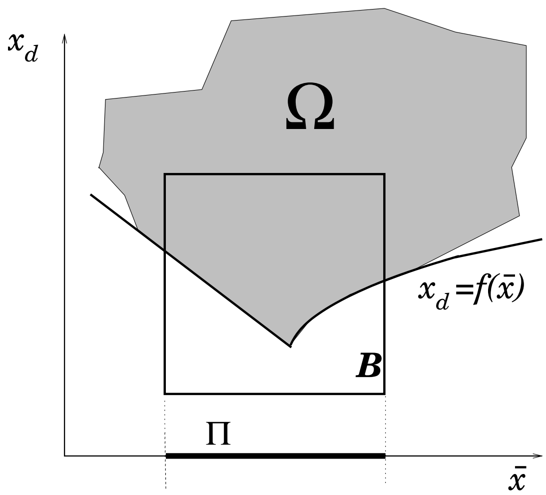


Fig 2. *A Lipschitz domain Ω and the localization by f in the ball B*

We denote by $\Pi = \{\bar{x}, x \in \partial\Omega \cap B\} \subset \mathbb{R}^{d-1}$ the projection of B onto the $(d-1)$ first components and by $\Pi_\lambda = \{\bar{x}, x \in \text{supp}\,(\lambda) \cap \Omega\}$. If ψ is a function defined on Σ, we denote by ψ_Σ the function defined on $[0,T) \times B \cap Q$ by $\psi_\Sigma(t,x) = \psi(t,\bar{x},f(\overline{x}))$. The function ψ_Σ does not depend on x_d and by an abuse of notation, we shall also denote by ψ_Σ the restriction of ψ_Σ to $[0,T) \times \Pi$. In the same way, if L_i is defined onto $[0,T] \times \Pi$, we also denote by L_i the function defined on $[0,T) \times B$ by $L_i(t,x) = L_i(t,\overline{x})$.

Let (ρ_n) be a sequence of mollifiers on $\mathbb{R}$ defined by

$$\rho_n(t) = n\rho(nt)$$

where ρ is a non-negative function of $\mathcal{C}_c^\infty(-1,0)$ such that $\displaystyle\int_{-1}^{0} \rho(t)\,dt = 1$. Let also R_n denote the function defined by

$$R_n(t) = \int_{-\infty}^{-t} \rho_n(s)\,ds\,.$$

For ε a positive number, we also denote by ρ_ε the function $t \mapsto \frac{1}{\varepsilon}\rho(\frac{t}{\varepsilon})$ and define the function $\omega_\varepsilon : \mathbb{R}^d \mapsto \mathbb{R}$ by

$$\omega_\varepsilon(x) = \int_{f(\overline{x})-x_d}^{0} \rho_\varepsilon(z)\,dz = \int_{\frac{f(\overline{x})-x_d}{\varepsilon}}^{0} \rho(z)\,dz.$$

Then

$$\nabla\omega_\varepsilon(x) = \rho_\varepsilon(f(\overline{x}) - x_d) \begin{pmatrix} -\nabla f(\overline{x}) \\ 1 \end{pmatrix}$$

On $\Omega \cap B$, the function ω_ε vanishes in a neighborhood of $\partial\Omega$ and equals 1 if $\text{dist}(x,\partial\Omega) > \varepsilon$. Also notice that, if $\psi \in H_{\text{div}}(\Omega)$ then

$$\int_\Omega \lambda\,\psi \cdot \nabla\omega_\varepsilon = -\int_\Omega \text{div}(\lambda\,\psi)\,\omega_\varepsilon \overset{\varepsilon\to 0}{\to} -\int_\Omega \text{div}(\lambda\,\psi) = -\int_{\partial\Omega} \lambda\,\psi\cdot\mathbf{n}\,.$$

Let $u \in L^\infty(Q \times (0,1))$ be an entropy process solution of Problem (3.1) and $\kappa \in \mathbb{R}$. We denote by $\mathcal{G}_x(t,x,u,\kappa)$ the quantity

$$\mathcal{G}_x(t,x,u,\kappa) = \Phi(t,x,u(t,x,\alpha),\kappa) - \nabla_x|\varphi(u)(t,x) - \varphi(\kappa)|\,. \tag{3.10}$$

For $w \in \mathbb{R}$, the function $\mathcal{F}_\varphi$ is defined by the formula

$$\mathcal{F}_\varphi(t,x,u,\kappa,w) = \mathcal{G}_x(t,x,u,\kappa) + \mathcal{G}_x(t,x,u,w) - \mathcal{G}_x(t,x,\kappa,w)\,. \tag{3.11}$$

3.4.2 First step in the proof of Theorem 3.4.1: study of the behavior of an entropy process solution near the boundary

We aim to prove the following proposition, which characterizes the way in which the boundary data $\overline{u}$ is assumed by the entropy process solution u (see [Ott96a] or [MPT00]).

Proposition 3.4.1 (Boundary entropy condition) *Let $u \in L^\infty(Q \times (0,1))$ be an entropy process solution of the problem (3.1) and let $\mathcal{F}_\varphi$ be defined by (3.11). Then, for all $\kappa \in \mathbb{R}$, for all nonnegative $\psi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^d)$, there holds:*

$$\limsup_{\varepsilon \to 0} \int_{Q^1} \mathcal{F}_\varphi(t,x,u(t,x,\alpha),\kappa,\overline{u}_\Sigma(t,x)) \cdot \nabla\omega_\varepsilon(x)\,\psi(t,x)\,\lambda(x)\,dx\,dt\,d\alpha \;\le 0 \tag{3.12}$$

In the case of a purely hyperbolic problem (that is to say $\varphi = 0$), Inequality (3.12) is the boundary condition written by Otto (see [Ott96a]). If the problem is strictly parabolic (that is to say $\varphi'(u) \ge \Phi_{min} > 0$), then Inequality (3.12) is trivially satisfied by any weak solution of the problem (see [MPT00]).

Lemma 3.4.1 *Let $\kappa \in \mathbb{R}$, let $\psi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^d)$ be such that $\psi \ge 0$ and $\mathrm{sgn}^+(\overline{u} - \kappa)\psi = 0$ on Σ. Then the following limit exists and is non-positive:*

$$\lim_{\varepsilon \to 0} \int_{Q^1} \left[\Phi^+(t,x,u(t,x,\alpha),\kappa) - \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^+\right] \cdot \nabla\omega_\varepsilon(x)\,\psi(t,x)\,\lambda(x)\,dx\,dt\,d\alpha \le 0\,. \tag{3.13}$$

Proof of Lemma 3.4.1. The function u being an entropy process solution, the linear application $\mathcal{I}^+$ defined on the vector field $\{\psi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^d)\,,\, \mathrm{sgn}^+(\overline{u} - \kappa)\psi = 0 \text{ on } \Sigma\,\}$ by

$$\mathcal{I}^+(\psi) = \left[\begin{array}{l} \displaystyle\iint_Q \int_0^1 \left(\eta_\kappa^+(u(t,x,\alpha))\,\psi_t(t,x) + \Phi_\kappa^+(t,x,u(t,x,\alpha)) \cdot \nabla\psi(t,x)\right)\,dx\,dtd\alpha \\ \displaystyle -\iint_Q \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^+ \cdot \nabla\psi(t,x)\,dx\,dt + \int_\Omega \eta_\kappa^+(u_0)\,\psi(0,x)\,dx \end{array}\right]$$

takes nonnegative values for nonnegative vectors ψ, so it is non-decreasing with respect to ψ. As $(\omega_\varepsilon)_\varepsilon$ is non-increasing when ε decrease to zero, $\mathcal{I}^+(\psi\,\lambda\,(1-\omega_\varepsilon))$ is non-decreasing and $\lim_{\varepsilon\to 0} \mathcal{I}^+(\psi\,\lambda\,(1-\omega_\varepsilon))$ exists, at least in $\overline{\mathbb{R}}$. Moreover, the sequence (ω_ε) converges to 1 everywhere on $\Omega \cap B$. Using the dominated convergence Theorem, we get

$$\lim_{\varepsilon \to 0} \int_{Q^1} \left[\Phi^+(t,x,u(t,x,\alpha),\kappa) - \nabla\left(\varphi(u)(t,x) - \varphi(\kappa)\right)^+\right] \cdot \nabla\omega_\varepsilon(x)\,\psi(t,x)\,\lambda(x)\,dx\,dt\,d\alpha \le 0,$$

and lemma 3.4.1 follows.$\square$

Remark 3.4.1 *The ε-limit considered in (3.13) only depends on the value of ψ on the boundary. This can be interesting to make the link between this lemma and the notion of entropy solution given in [MPT00]. Moreover, it is useful in the final part of the proof.*

To prove this fact, let ψ be a nonnegative function of $\mathcal{C}_c^\infty([0,T)\times\mathbb{R}^d)$ such that $\mathrm{sgn}^+(\overline{u}-\kappa)\psi=0$ on Σ. There exists a constant C_ψ such that $|\psi(t,x)-\psi_\Sigma(t,x)|\leq C_\psi\,|x_d-f(\overline{x})|$ and, denoting by ϑ the changing of variable defined by

$$\vartheta(x)=(\bar{x},x_d+f(\bar{x}))$$

and by $\widetilde{u}$ the function $(t,x,\alpha)\mapsto u(t,\vartheta(x),\alpha)$ we have:

$$\begin{aligned}
&\Big|\int_{Q^1}\Big[\Phi^+(t,x,u(t,x),\kappa)-\nabla\left(\varphi(u(t,x))-\varphi(\kappa)\right)^+\Big]\cdot\nabla\omega_\varepsilon(x)\left(\psi(t,x)-\psi_\Sigma(t,x)\right)\lambda(x)\,dx\,dt\,d\alpha\Big|\\
\leq\;& C_\psi\iint_{t,\overline{x}}\int_{x_d=f(\overline{x})}^{f(\overline{x})+\varepsilon}\int_0^1\Big[|\Phi_d^+(t,x,u,\kappa)|+|\nabla\left(\varphi(u)-\varphi(\kappa)\right)^+|\Big]|\nabla\omega_\varepsilon|\,|x_d-f(\overline{x})|\,\lambda\,dx_d\,d\overline{x}\,dt\,d\alpha\\
=\;& C_\psi\iint_{t,\overline{x}}\int_{\widetilde{x_d}=0}^{\varepsilon}\int_0^1\Big[|\Phi_d^+(t,x,\widetilde{u},\kappa)|+|\nabla\left(\varphi(\widetilde{u})-\varphi(\kappa)\right)^+|\Big]\sqrt{1+|\nabla f(\overline{x})|^2}\,\rho_\varepsilon(-\widetilde{x_d})\,\widetilde{\lambda}\,\widetilde{x_d}\,d\widetilde{x_d}\,d\overline{x}\,dt\,d\alpha\\
\leq\;& C(\psi,T,|B|)\Big(||\Phi^+(\cdot,\cdot,u,\kappa)||_{L^2(Q^1)}+||\nabla(\varphi(u)-\varphi(\kappa))||_{L^2(Q)}\Big)\Big(\int_{\widetilde{x_d}=0}^{\varepsilon}\rho_\varepsilon(-\widetilde{x_d})^2\,\widetilde{x_d}^2\,d\widetilde{x_d}\Big)^{1/2}\\
\leq\;& C(\psi,T,|B|)\,\sqrt{\varepsilon}\,,
\end{aligned}$$

where C is a constant independent of ε.
The same results hold when the entropy function $u\mapsto(u-\kappa)^-$ is considered, in that case the function ψ has to satisfy the condition $\mathrm{sgn}^-(\overline{u}-\kappa)\,\psi|_\Sigma=0$.

Now, we are able to prove Proposition 3.4.1.

PROOF OF PROPOSITION 3.4.1.
By definition, $\Pi_\lambda\subset\Pi$ and $d_{max}=dist(\Pi_\lambda,\Pi^c)>0$. So if we define Π_λ^δ by $\Pi_\lambda^\delta=\Pi_\lambda+B(0,\delta)$, for $\delta<d_{max}$, we have $\Pi_\lambda\subset\Pi_\lambda^\delta\subset\Pi$. Let $\theta^\delta:\mathbb{R}^{d-1}\mapsto[0,1]$ be a $\mathcal{C}_c^\infty$ function with support included in Π_λ^δ and such that $\theta^\delta=1$ on Π_λ (θ^δ approximates the characteristic function of the set Π_λ)

Fix η positive, δ such that $0<\delta<d_{max}$ and $\psi\in\mathcal{C}_c^\infty([0,T)\times\mathbb{R}^d)$ such that $\psi\geq0$. Let T_ψ be such that $\psi=0$ on $[T\psi,T)$. As the function $\overline{u}_\Sigma$ is continuous on the closure of $[0,T_\psi]\times\Pi_\lambda^\delta$, a compact subset of $\mathbb{R}_+\times\mathbb{R}^{d-1}$, there exists α a positive number such that

$$|\overline{u}_\Sigma(t,\bar{x})-\overline{u}_\Sigma(s,\overline{y})|<\eta$$

whenever $(t,\overline{x})$, $(s,\overline{y})\in[0,T_\psi]\times\Pi_\lambda$ and $|(t,\overline{x})-(s,\overline{y})|\leq\alpha$. Moreover, there exists $P_\eta\in\mathbb{N}$ and some balls $V_1^\eta,\ldots,V_{P_\eta}^\eta$ of diameter less than α such that

$$[0,T_\psi]\times\Pi_\lambda\subset\bigcup_{i=1}^{P_\eta}V_i^\eta.$$

Let $(L_i^\eta) : \mathbb{R}_+ \times \mathbb{R}^{d-1} \mapsto \mathbb{R}$ be a $\mathcal{C}^\infty$ partition of unity on $[\,0, T_\psi] \times \Pi_\lambda$ subordinate to the open cover (V_i^η) with supp $(L_i^\eta) \subset \mathbb{R}_+ \times \Pi_\lambda^\delta$. For $i \in \{1, \ldots, P_\eta\}$, pick up $(t_i, \overline{x}_i) \in V_i^\eta \cap [0, T_\psi] \times \Pi_\lambda$ and set

$$w_i^\eta = \overline{u}_\Sigma(t_i, \overline{x}_i) + \eta$$

in order to ensure $\overline{u}_\Sigma < w_i^\eta < \overline{u}_\Sigma + 2\eta$ on supp L_i^η.
Then the function $\Psi_i : (t,x) \mapsto \psi(t,x)\, L_i^\eta(t,\overline{x})\theta^\delta(\overline{x})$ satisfies the condition

$$\operatorname{sgn}^+(\overline{u} - \kappa \top w_i^\eta)\, \Psi_i = 0 \ \text{ on } \Sigma$$

and from Lemma 3.4.1 and the fact that $\theta^\delta \lambda = \lambda$ is deduced:

$$\lim_{\varepsilon \to 0} \int_{Q^1} \left[\Phi^+(t,x,u,\kappa \top w_i^\eta) - \nabla\left(\varphi(u) - \varphi(\kappa \top w_i^\eta)\right)^+\right] \cdot \nabla \omega_\varepsilon\, \psi \lambda\, L_i^\eta\, dx\, dt\, d\alpha \ \leq 0\,. \tag{3.14}$$

Now, write

$$\begin{aligned}
& \int_{Q^1} \left[\Phi^+(t,x,u,\kappa \top \overline{u}) - \nabla\left(\varphi(u) - \varphi(\kappa \top \overline{u})\right)^+\right] \cdot \nabla \omega_\varepsilon\, \psi \lambda\, L_i^\eta\, dx\, dt\, d\alpha \\
= & \int_{Q^1} \left[\Phi^+(t,x,u,\kappa \top w_i^\eta) - \nabla\left(\varphi(u) - \varphi(\kappa \top w_i^\eta)\right)^+\right] \cdot \nabla \omega_\varepsilon\, \psi \lambda\, L_i^\eta\, dx\, dt\, d\alpha \\
+ & \int_{Q^1} \left[\Phi^+(t,x,u,\kappa \top \overline{u}_\Sigma) - \Phi^+(t,x,u,\kappa \top w_i^\eta)\right] \cdot \nabla \omega_\varepsilon\, \psi \lambda\, L_i^\eta\, dx\, dt\, d\alpha \\
- & \int_{Q^1} \nabla\Big[\left(\varphi(u) - \varphi(\kappa \top \overline{u}_\Sigma)\right)^+ - \left(\varphi(u) - \varphi(\kappa \top w_i^\eta)\right)^+\Big] \cdot \nabla \omega_\varepsilon\, \psi \lambda\, L_i^\eta\, dx\, dt\, d\alpha\,.
\end{aligned}$$

Using the Lipschitz continuity of $\kappa \mapsto \Phi(t,x,u(t,x,\alpha),\kappa)$ (which is uniform with respect to $(t,x,\alpha) \in Q \times (\,0,1))$ and the fact that $\int_{\mathbb{R}} |\nabla \omega_\varepsilon(t,\overline{x},x_d)| dx_d$ is bounded independently of ε, the second term of the right hand-side of this equality can be estimated independently of ε:

$$\begin{aligned}
& \Big| \int_{Q^1} \left[\Phi^+(t,x,u,\kappa \top \overline{u}_\Sigma) - \Phi^+(t,x,u,\kappa \top w_i^\eta)\right] \cdot \nabla \omega_\varepsilon\, \psi \lambda\, L_i^\eta\, dx\, dt\, d\alpha \Big| \\
\leq \ & \mathcal{L}\, ||\psi||_\infty \iint_{[0,T_\psi] \times \Pi_\lambda} |\overline{u}_\Sigma(t,\overline{x}) - w_i^\eta|\, L_i^\eta(t,\overline{x})\, d\overline{x}\, dt \\
\leq \ & 2\,\mathcal{L}\, ||\psi||_\infty\, \eta \iint_{[0,T_\psi] \times \Pi_\lambda} L_i^\eta(t,\overline{x})\, d\overline{x}\, dt\,.
\end{aligned}$$

Therefore, from (3.14) is deduced

$$\begin{aligned}
& \limsup_{\varepsilon \to 0} \int_{Q^1} \left[\Phi^+(t,x,u,\kappa \top \overline{u}_\Sigma) - \nabla\left(\varphi(u) - \varphi(\kappa \top \overline{u}_\Sigma)\right)^+\right] \cdot \nabla \omega_\varepsilon\, \psi \lambda L_i^\eta\, dx\, dt\, d\alpha \\
\leq \ & 2\,\mathcal{L}\, ||\psi||_\infty\, \eta \iint_{[0,T_\psi] \times \Pi_\lambda} L_i^\eta(t,\overline{x})\, d\overline{x}\, dt \\
& - \liminf_{\varepsilon \to 0} \int_{Q^1} \nabla\Big[\left(\varphi(u) - \varphi(\kappa \top \overline{u}_\Sigma)\right)^+ - \left(\varphi(u) - \varphi(\kappa \top w_i^\eta)\right)^+\Big] \cdot \nabla \omega_\varepsilon\, \psi \lambda\ L_i^\eta\, dx\, dt\,.
\end{aligned}$$

The second term of the right-hand side of this inequality writes $\liminf\limits_{\varepsilon\to 0}\displaystyle\iint_Q \nabla W\cdot\nabla\omega_\varepsilon\,\psi\lambda\,L_i^\eta\,dx\,dt$ where the function W is defined by

$$W(t,x)=\left[\,(\varphi(u(t,x))-\varphi(\kappa\top\overline{u}_\Sigma))^+-(\varphi(u(t,x))-\varphi(\kappa\top w_i^\eta))^+\,\right]$$

is an element of $L^2(\,0,T;H_0^1(\Omega))$ and it is nonnegative on supp $(\psi\lambda L_i^\eta)$ and $\psi\lambda L_i^\eta$ is a regular nonnegative function. Hence, the infimum limit considered above is non-negative and

$$\begin{aligned}\limsup_{\varepsilon\to 0}\int_{Q^1}\left[\Phi^+(t,x,u,\kappa\top\overline{u}_\Sigma)-\nabla\left(\varphi(u)-\varphi(\kappa\top\overline{u}_\Sigma)\right)^+\right]\cdot\nabla\omega_\varepsilon\,\psi\lambda\,L_i^\eta\,dx\,dt\,d\alpha&\\ \le\,2\,\mathcal{L}\,||\psi||_\infty\,\eta\iint_{[0,T_\psi]\times\Pi_\lambda}L_i^\eta(t,\overline{x})\,d\overline{x}\,dt&\end{aligned}$$

As (L_i^η) is a partition of unity, the equality $\psi\lambda\displaystyle\sum_{i=1}^{P_\eta}L_i^\eta=\psi\lambda$ holds and, summing the above result for $i\in\{1,\dots,P_\eta\}$, we get the estimate:

$$\limsup_{\varepsilon\to 0}\int_{Q^1}\left[\Phi^+(t,x,u,\kappa\top\overline{u}_\Sigma)-\nabla\left(\varphi(u)-\varphi(\kappa\top\overline{u}_\Sigma)\right)^+\right]\cdot\nabla\omega_\varepsilon\,\psi\lambda\,dx\,dt\,d\alpha\le\,2\,\mathcal{L}\,||\psi||_\infty\,\eta\,T_\psi|\Pi_\lambda|\,.$$

The left hand-side is independent of η, and since it can be taken as small as we want, the supremum limit considered is non-positive:

$$\limsup_{\varepsilon\to 0}\int_{Q^1}\left[\Phi^+(t,x,u,\kappa\top\overline{u}_\Sigma)-\nabla\left(\varphi(u)-\varphi(\kappa\top\overline{u}_\Sigma)\right)^+\right]\cdot\nabla\omega_\varepsilon\,\psi\lambda\,dx\,dt\,d\alpha\le 0\,.$$

The same kind of study can be made to prove the following inequality:

$$\limsup_{\varepsilon\to 0}\int_{Q^1}\left[\Phi^-(t,x,u,\kappa\bot\overline{u}_\Sigma)-\nabla\left(\varphi(u)-\varphi(\kappa\bot\overline{u}_\Sigma)\right)^-\right]\cdot\nabla\omega_\varepsilon\,\psi\lambda\,dx\,dt\,d\alpha\le 0\,.$$

Then one concludes the proof of Proposition 3.4.1 by using the formula

$$\begin{aligned}\mathcal{F}_\varphi(t,x,u,\kappa,w)\;&=\;[\Phi^+(t,x,u,\kappa\top w)-(\varphi(u)-\varphi(\kappa\top w))^+]\\&\quad+[\Phi^-(t,x,u,\kappa\bot w)-(\varphi(u)-\varphi(\kappa\bot w))^-]\,.\end{aligned}$$

□

3.4.3 Second step in the proof of Theorem 3.4.1: inner comparison

Let u and $v\in L^\infty(Q\times(\,0,1))$ be two entropy process solutions of Problem (3.1). The following result of comparison between u and v involving test functions which *vanish* on the boundary of Ω can be proved (see [Car99] or [EGHM00]):

Proposition 3.4.2 (Inner comparison): *Let u and $v \in L^\infty(Q\times(0,1))$ be two entropy process solutions of Problem (3.1). Let ξ be a non-negative function of $\mathcal{C}^\infty([0,T)\times\mathbb{R}^d\times[0,T)\times\mathbb{R}^d)$ such that:*

$$\left\{\begin{array}{l} \textit{for all } (s,y)\in Q\,,\ (t,x)\longmapsto \xi(t,x,s,y)\in\mathcal{C}_c^\infty([0,T)\times\Omega)\,,\\ \textit{for all } (t,x)\in Q\,,\ (s,y)\longmapsto \xi(t,x,s,y)\in\mathcal{C}_c^\infty([0,T)\times\Omega)\,. \end{array}\right.$$

Then the following inequality holds:

$$\begin{aligned}&\int_{Q^1}\int_{Q^1}\left[\begin{array}{c} |u(t,x,\alpha)-v(s,y,\beta)|(\xi_t+\xi_s)\\ +\mathcal{G}_x(t,x,u(t,x,\alpha),v(s,y,\beta))\cdot\nabla_x\xi+\mathcal{G}_y(s,y,v(s,y,\beta),u(t,x,\alpha))\cdot\nabla_y\xi\\ -\nabla_x|\varphi(u)(t,x)-\varphi(v)(s,y)|\cdot\nabla_y\xi-\nabla_y|\varphi(u)(t,x)-\varphi(v)(s,y)|\cdot\nabla_x\xi \end{array}\right]dx\,dt\,d\alpha\,dy\,ds\,d\beta\\ &+\int_{Q^1}\int_\Omega |u_0(x)-v(s,y,\beta)|\,\xi(0,x,s,y)\,dx\,dy\,ds\,d\beta\\ &+\int_{Q^1}\int_\Omega |u_0(y)-u(t,x,\alpha)|\,\xi(t,x,0,y)\,dx\,dt\,d\alpha\,dy\ \geq 0\,.\end{aligned}\tag{3.15}$$

3.4.4 Third step in the proof of Theorem 3.4.1: the doubling variable method

An entropy process solution u of Problem (3.1) completely assumes the initial condition at time $t=0$, for example meaning that

$$\operatorname*{ess\,lim}_{t\to 0^+}\int_\Omega\int_0^1 |u(t,x,\alpha)-u_0(x)|\,dx\,d\alpha = 0\,.$$

As far as the boundary condition is concerned, this is not the case for the function φ may degenerate. That is why the work on time variables is distinguished from the work on space variables.

Doubling time variables

Let $\zeta\in\mathcal{C}_c^\infty([0,T)\times\mathbb{R}^d\times\mathbb{R}^d)$ be a non-negative function such that

$$\left\{\begin{array}{l} \text{for all } (t,y)\in Q\,,\ x\longmapsto\zeta(t,x,y)\in\mathcal{C}_c^\infty(\Omega)\,,\\ \text{for all } (t,x)\in Q\,,\ y\longmapsto\zeta(t,x,y)\in\mathcal{C}_c^\infty(\Omega)\,. \end{array}\right.$$

Setting $\xi(t,x,s,y)=\rho_n(t-s)\zeta(t,x,y)$ in Inequality (3.15), and noticing that $\xi(t,x,0,y)=0$ one gets:

$$\begin{aligned}&\int_{Q^1}\int_{Q^1}\left[\begin{array}{c} |u(t,x,\alpha)-v(s,y,\beta)|\zeta_t\,\rho_n(t-s)\\ +\mathcal{G}_x(t,x,u(t,x,\alpha),v(s,y,\beta))\cdot\nabla_x\zeta\,\rho_n(t-s)\\ +\mathcal{G}_y(s,y,v(s,y,\beta),u(t,x,\alpha))\cdot\nabla_y\zeta\,\rho_n(t-s)\\ -\nabla_x|\varphi(u)(t,x)-\varphi(v)(s,y)|\cdot\nabla_y\xi\,\rho_n(t-s)\\ -\nabla_y|\varphi(u)(t,x)-\varphi(v)(s,y)|\cdot\nabla_x\xi\,\rho_n(t-s) \end{array}\right]dx\,dt\,d\alpha\,dy\,ds\,d\beta\\ &+\int_{Q^1}\int_\Omega |u_0(x)-v(s,y,\beta)|\,\zeta(0,x,y)\,\rho_n(-s)\,dx\,dy\,ds\,d\beta\geq 0\,.\end{aligned}\tag{3.16}$$

The theorem of convergence in means allows to calculate the limit as $n \to +\infty$ of the first term of this inequality. To estimate the behavior of the second term, set $\kappa = u_0(x)$ and $\psi(s,y) = R_n(s)\,\zeta(0,x,y)$ in the following entropy inequality satisfied by the function v:

$$\begin{aligned}&\int_{Q^1} |v(s,y,\beta)-\kappa|\,\psi_t + (\Phi(s,y,v(s,y,\beta),\kappa) - \nabla|\varphi(v)(s,y)-\varphi(\kappa)|)\cdot\nabla\psi\,dy\,ds\,d\beta\\ &+\int_\Omega |u_0(y)-\kappa|\psi(0,y)\,dy \geq 0\,.\end{aligned}$$

Integrating the result w.r.t. x yields an upper bound of the second term of the inequality (3.16), that is:

$$\begin{aligned}&\int_{Q^1}\int_\Omega |u_0(x)-v(s,y,\beta)|\,\zeta(0,x,y)\,\rho_n(-s)\,dx\,dy\,ds\,d\beta\\ \leq\ &\int_{Q^1}\int_\Omega (\Phi(s,y,v(s,y,\beta),u_0(x)) - \nabla|\varphi(v)(s,y)-\varphi(u_0(x))|)\cdot\nabla_y\zeta\,R_n(s)\,dx\,dy\,ds\,d\beta\\ +\ &\int_\Omega\int_\Omega |u_0(x)-u_0(y)|\zeta(0,x,y)\,dx\,dy\,.\end{aligned}$$

As (R_n) is a bounded sequence of functions converging to zero, the limit of the first term of the right hand-side of the previous inequality is zero and there holds the following estimate, which, we recall, signifies that the entropy process solution v completely assumes the initial condition u_0:

$$\begin{aligned}&\limsup_{n\to+\infty}\int_{Q^1}\int_\Omega |u_0(x)-v(s,y,\beta)|\,\zeta(0,x,y)\,\rho_n(-s)\,dx\,dy\,ds\,d\beta\\ \leq\ &\int_\Omega\int_\Omega |u_0(x)-u_0(y)|\zeta(0,x,y)\,dx\,dy\,.\end{aligned}$$

Eventually, the following proposition holds:

Proposition 3.4.3 *Let u and $v \in L^\infty(Q\times(0,1)$ be two entropy process solution of the problem (3.1). Let $\zeta \in \mathcal{C}_c^\infty([0,T)\times\mathbb{R}^d\times\mathbb{R}^d)$ be a non-negative function such that*

$$\left\{\begin{array}{l}\text{for all } (t,y)\in Q\,,\ x\longmapsto\zeta(t,x,y)\in\mathcal{C}_c^\infty(\Omega)\,,\\ \text{for all } (t,x)\in Q\,,\ y\longmapsto\zeta(t,x,y)\in\mathcal{C}_c^\infty(\Omega)\,.\end{array}\right. \tag{3.17}$$

Then the following partial result of comparison holds:

$$\begin{aligned}&\int_{Q^1}\int_{\Omega^1}\left[\begin{array}{c}|u(t,x,\alpha)-v(t,y,\beta)|\zeta_t\\ +\mathcal{G}_x(t,x,u(t,x,\alpha),v(t,y,\beta))\cdot\nabla_x\zeta+\mathcal{G}_y(t,y,v(t,y,\beta),u(t,x,\alpha))\cdot\nabla_y\zeta\\ -\nabla_x|\varphi(u)(t,x)-\varphi(v)(t,y)|\cdot\nabla_y\zeta\ -\nabla_y|\varphi(u)(t,x)-\varphi(v)(t,y)|\cdot\nabla_x\zeta\end{array}\right]dx\,dt\,d\alpha\,dy\,d\beta\\ &+\int_\Omega\int_\Omega |u_0(x)-u_0(y)|\,\zeta(0,x,y)\,dx\,dy\ \geq 0\,.\end{aligned} \tag{3.18}$$

Doubling space variables

We now follow the lines of the proof of uniqueness given by Mascia, Porretta and Terracina [MPT00]. First, we would like to consider test-functions which do not necessarily vanish on $\partial\Omega$ and are localized into the ball B.
For $\varepsilon > 0$ define ξ to be the function

$$\xi(t,x,y) = \psi(t,x)\,\rho_m(\overline{x}-\overline{y})\,\rho_n(x_d-y_d).$$

Using the properties of ρ we have

$$\begin{array}{l} \text{for all } (t,x)\in Q\,,\ y\longmapsto \xi(t,x,y)\in\mathcal{C}_c^\infty(\Omega)\,,\\ \text{for all } (t,x)\in[\,0,T)\times\operatorname{supp}(\lambda)\,,\ \operatorname{supp}_y\xi(t,x,\cdot)\subset B\,. \end{array} \tag{3.19}$$

For $\varepsilon > 0$ define ζ to be the function

$$\zeta(t,x,y)\longmapsto\ \omega_\varepsilon(x)\,\xi(t,x,y)\,\lambda(x)$$

(recall that ω_ε is defined by (3.10)). Then, Assumption (3.17) of Proposition 3.4.3 holds and, with this particular choice of function ζ, Inequality (3.18) turns into the inequality

$$\begin{aligned}
&\int_{Q^1}\int_{\Omega^1}\left[\begin{array}{c} |u-\widehat{v}|\,\omega_\varepsilon(x)\,(\xi\lambda)_t\\ +\Big(\mathcal{G}_x(t,x,u,\widehat{v})\cdot\nabla_x(\xi\,\lambda)+\mathcal{G}_y(t,y,\widehat{v},u)\cdot\nabla_y(\xi\,\lambda\Big)\,\omega_\varepsilon(x)\\ -\Big(\nabla_x|\varphi(u)-\varphi(\widehat{v})|\cdot\nabla_y(\xi\,\lambda)+\nabla_y|\varphi(u)-\varphi(\widehat{v})|\cdot\nabla_x(\xi\,\lambda)\Big)\,\omega_\varepsilon(x)\end{array}\right]dx\,dt\,d\alpha\,dy\,d\beta\\
&+\int_{Q^1}\int_{\Omega^1}\mathcal{G}_x(t,x,u,\widehat{v})\cdot\nabla\omega_\varepsilon(x)\ \xi\,\lambda\,dx\,dt\,d\alpha\,dy\,d\beta\\
&-\int_{Q^1}\int_{\Omega^1}\nabla_y|\varphi(u)(t,x)-\varphi(v)(t,y)|\cdot\nabla\omega_\varepsilon(x)\,\xi\,\lambda\,dx\,dt\,d\alpha\,dy\,d\beta\\
&+\int_\Omega\int_\Omega|u_0(x)-u_0(y)|\,(\xi\,\lambda)(0,x,y)\,\omega_\varepsilon(x)\,dx\,dy\ \geq 0\,,
\end{aligned}$$

where

$$u=u(t,x,\alpha)\ \text{ and }\ \widehat{v}=v(t,y,\beta)\,.$$

Using Formula (3.11), this rewrites:

$$\begin{aligned}
&\int_{Q^1}\int_{\Omega^1}\left[\begin{array}{c} |u-\widehat{v}|\,\omega_\varepsilon(x)\,(\xi\,\lambda)_t\\ +(\mathcal{G}_x(t,x,u,\widehat{v})\cdot\nabla_x(\xi\,\lambda)+\mathcal{G}_y(t,y,\widehat{v},u)\cdot\nabla_y(\xi\,\lambda)\,\omega_\varepsilon(x)\\ -(\nabla_x|\varphi(u)-\varphi(\widehat{v})|\cdot\nabla_y(\xi\,\lambda)+\nabla_y|\varphi(u)-\varphi(\widehat{v})|\cdot\nabla_x(\xi\,\lambda))\,\omega_\varepsilon(x)\end{array}\right]dx\,dt\,d\alpha\,dy\,d\beta\\
&+\int_{Q^1}\int_{\Omega^1}\mathcal{F}_\varphi(t,x,u,\widehat{v},\overline{u}_\Sigma)\cdot\nabla\omega_\varepsilon(x)\,\xi\,\lambda\,dx\,dt\,d\alpha\,dy\,d\beta\\
&+\int_\Omega\int_\Omega|u_0(x)-u_0(y)|\,(\xi\,\lambda)(0,x,y)\,\omega_\varepsilon(x)\,dx\,dy\ \geq\ A+B+C\,,
\end{aligned} \tag{3.20}$$

where:

$$\begin{aligned}
A &= \int_Q \int_\Omega \nabla_y |\varphi(u)(t,x) - \varphi(v)(t,y)| \cdot \nabla \omega_\varepsilon(x)\, \xi\, \lambda \, dx\, dt, dy\,,\\
B &= -\int_{Q^1} \int_{\Omega^1} \mathcal{G}_x(t,x,\widehat{v},\overline{u}_\Sigma) \cdot \nabla \omega_\varepsilon(x)\, \xi\, \lambda \, dx\, dt\, d\alpha\, dy\, d\beta\,,\\
C &= \int_{Q^1} \int_{\Omega^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma) \cdot \nabla \omega_\varepsilon(x)\, \xi\, \lambda \, dx\, dt\, d\alpha\, dy\, d\beta\,.
\end{aligned}$$

Using Proposition 3.4.1 and taking the supremum limit of both hand-sides of the previous inequality with respect to ε yields the following result:

$$\begin{aligned}
&\int_{Q^1} \int_{\Omega^1} \left[\begin{array}{c} |u - \widehat{v}|\, (\xi\, \lambda)_t \\ +\mathcal{G}_x(t,x,u,\widehat{v}) \cdot \nabla_x(\xi\, \lambda) + \mathcal{G}_y(t,y,\widehat{v},u) \cdot \nabla_y(\xi\, \lambda) \\ -(\nabla_x |\varphi(u) - \varphi(\widehat{v})| \cdot \nabla_y(\xi\, \lambda) + \nabla_y |\varphi(u) - \varphi(\widehat{v})| \cdot \nabla_x(\xi\, \lambda) \end{array} \right] dx\, dt\, d\alpha\, dy\, d\beta \\
&+ \int_\Omega \int_\Omega |u_0(x) - u_0(y)|\, (\xi\, \lambda)(0,x,y)\, dx\, dy \;\geq\; \limsup_{\varepsilon \to 0}\, (A + B + C\,)\,,
\end{aligned}$$

or, as well (using Formula (3.10)):

$$\begin{aligned}
&\int_{Q^1} \int_{\Omega^1} \left[\begin{array}{c} |u - \widehat{v}|\, (\xi\, \lambda)_t \\ +\Phi(t,x,u,\widehat{v}) \cdot \nabla_x(\xi\, \lambda) + \Phi(t,y,\widehat{v},u) \cdot \nabla_y(\xi\, \lambda) \\ -(\nabla_x |\varphi(u) - \varphi(\widehat{v})| + \nabla_y |\varphi(u) - \varphi(\widehat{v})|) \cdot (\nabla_y + \nabla_x)(\xi\, \lambda) \end{array} \right] dx\, dt\, d\alpha\, dy\, d\beta \\
&+ \int_\Omega \int_\Omega |u_0(x) - u_0(y)|\, (\xi\, \lambda)(0,x,y)\, dx\, dy \;\geq\; \limsup_{\varepsilon \to 0}\, (A + B + C\,)\,,
\end{aligned}$$

The study of the behavior of A, B and C as $[\varepsilon \to 0]$ and the doubling variable technique itself interfere with each other. We use the definition of ξ.

$$\xi(t,x,y) = \psi(t,x)\, \rho_m(\overline{x} - \overline{y})\, \rho_n(x_d - y_d)\,.$$

Notice that, for n large enough, and m sufficiently large compared with n, Assumption (3.19) holds and that, besides, C does not depend on m and n:

$$C = \int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma) \cdot \nabla \omega_\varepsilon(x)\, \psi\, \lambda \, dx\, dt\, d\alpha\,.$$

Moreover, Inequality (3.21) rewrites:

$$\begin{aligned}
&\int_{Q^1} \int_{\Omega^1} \left[\begin{array}{c} |u - \widehat{v}|\, \rho_m\, \rho_n (\psi\, \lambda)_t \\ +\Phi(t,x,u,\widehat{v}) \cdot \nabla_x(\psi\, \lambda) \rho_m\, \rho_n \\ -(\nabla_x |\varphi(u) - \varphi(\widehat{v})| + \nabla_y |\varphi(u) - \varphi(\widehat{v})|) \cdot \nabla_x(\psi\, \lambda)\, \rho_m\, \rho_n \end{array} \right] dx\, dt\, d\alpha\, dy\, d\beta \\
&+ \int_\Omega \int_\Omega |u_0(x) - u_0(y)|\, (\psi\, \lambda)(0,x)\, \rho_m\, \rho_n\, dx\, dy \;\geq\; \limsup_{\varepsilon \to 0}\, (A + B + C\,) + D\,,
\end{aligned} \tag{3.21}$$

where

$$D = -\int_{Q^1} \int_{\Omega^1} [\Phi(t,x,u,\widehat{v}) - \Phi(t,y,u,\widehat{v})] \cdot \nabla_x(\rho_m\, \rho_n)\, \psi\, \lambda\,.$$

If the flux-function F does not depend on the (t,x)-variables, then $D=0$. More generally, one can prove (see [CH99]):

$$D \geq E = -C(F,\psi)\sup\left\{\int_{Q^1} |v(t,\overline{x},x_d,\alpha) - v(t,\overline{x}+\overline{h},x_d+k,\alpha)|dt\,d\overline{x}\,dx_d\,d\alpha\,,\ |\overline{h}| \leq \frac{1}{m}\,,\ |k| \leq \frac{1}{n}\right\}.$$

So $\lim\limits_{m,n\to+\infty} D = 0$

Going back to the study of A, B, we write $A+B = I + J^y + J^x$ where

$$\begin{aligned}
I &= -\int_{Q^1}\int_{\Omega^1} (\Phi(t,x,\widehat{v},\overline{u}_\Sigma(t,\overline{x}))\cdot\nabla\omega_\varepsilon(x)\,\xi\,\lambda\,dx\,dt\,d\alpha\,dy\,d\beta\,,\\
J^y &= \int_Q\int_\Omega \nabla_y|\varphi(u)(t,x) - \varphi(v)(t,y)|\cdot\nabla\omega_\varepsilon(x)\,\xi\,\lambda\,dx\,dt\,dy\,,\\
J^x &= \int_Q\int_\Omega \nabla_x|\varphi(\widehat{v}) - \varphi(\overline{u}_\Sigma(t,\overline{x}))|\cdot\nabla\omega_\varepsilon(x)\,\xi\,\lambda\,dx\,dt\,dy\,.
\end{aligned}$$

Recall that

$$\nabla\omega_\varepsilon(x) = \rho_\varepsilon(f(\overline{x}) - x_d)\begin{pmatrix}-\nabla f(\overline{x})\\ 1\end{pmatrix},$$

so that

$$\widetilde{I} = \lim_{\varepsilon\to 0} I = -\int_{Q^1}\int_{\Pi\times(0,1)} (\Phi(t,\overline{x},f(\overline{x}),\widehat{v},\overline{u}_\Sigma(t,\overline{x}))\cdot\begin{pmatrix}-\nabla f(\overline{x})\\ 1\end{pmatrix}(\xi\,\lambda)_{\Sigma_x}\,d\overline{x}\,d\alpha\ dt\,dy\,d\beta\,.$$

where the index Σ_x signifies that the transformation only concerns the x variable. Here for example $(\xi\,\lambda)_{\Sigma_x}(t,x,y) = \xi(t,\overline{x},f(\overline{x}),y)\lambda(\overline{x},f(\overline{x}))$. To study J^x, we notice that the function $\overline{u}_\Sigma$ does not depend on x_d, so that:

$$\widetilde{J^x} = \lim_{\varepsilon\to 0} J^x = -\int_{[0,T)\times\Pi}\int_\Omega \nabla_{\overline{x}}\,|\varphi(\widehat{v}) - \varphi(\overline{u}_\Sigma(t,\overline{x}))|\cdot\nabla f(\overline{x})\,(\xi\,\lambda)_{\Sigma_x}\,d\overline{x}\,dt\,dy\,.$$

Integrating by parts with respect to $\overline{x}$ in $\widetilde{J^x}$ yields: $\widetilde{J^x} = \widetilde{J^x_f} + \widetilde{J^x_\psi} + \widetilde{J^x_{\rho_m}} + \widetilde{J^x_{\rho_n}}$, where:

$$\widetilde{J^x_f} = \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\widehat{v}) - \varphi(\overline{u}_\Sigma)|\Delta f(\overline{x})\,(\psi\,\lambda)_{\Sigma_x}\,\rho_m(\overline{x}-\overline{y})\,\rho_n(f(\overline{x}) - y_d)\,d\overline{x}\,dt\,dy\,,$$

$$\widetilde{J^x_\psi} = \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\widehat{v}) - \varphi(\overline{u}_\Sigma)|(\nabla f(\overline{x})\cdot\Big(\nabla_{\overline{x}}\,((\psi\,\lambda)_{\Sigma_x})\rho_m(\overline{x}-\overline{y})\,\rho_n(f(\overline{x}) - y_d)\,d\overline{x}\,dt\,dy\,,$$

$$\widetilde{J^x_{\rho_m}} = \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\widehat{v}) - \varphi(\overline{u}_\Sigma)|\nabla f(\overline{x})\cdot\nabla_{\overline{x}}\,\rho_m(\overline{x}-\overline{y})\,\rho_n(f(\overline{x}) - y_d)\,\psi\,\lambda\,d\overline{x}\,dt\,dy\,,$$

$$\widetilde{J^x_{\rho_n}} = \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\widehat{v}) - \varphi(\overline{u}_\Sigma)|\,|\nabla f(\overline{x})|^2\,\rho_m(\overline{x}-\overline{y})\,\rho'_n(f(\overline{x}) - y_d)\,(\psi\,\lambda)_{\Sigma_x}\,d\overline{x}\,dt\,dy\,.$$

On the other hand, integrating by parts in J^y with respect to y, recall that the boundary condition $\varphi(u) = \varphi(\overline{u})$ on Σ is strongly satisfied according to Definition 3.3.1, we get:

$$\widetilde{J^y} = \lim_{\varepsilon\to 0} J^y = -\int_{[0,T)\times\Pi}\int_\Omega |\varphi(\overline{u}_\Sigma(t,\overline{x})) - \varphi(\widehat{v})| \begin{pmatrix} -\nabla f(\overline{x}) \\ 1 \end{pmatrix} \cdot \nabla_y(\xi\,\lambda)(t,\overline{x},f(\overline{x}),y)\,d\overline{x}\,dt\,dy\,,$$

and, developing the scalar product:

$$\begin{aligned}
\widetilde{J^y} &= \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\overline{u}_\Sigma(t,\overline{x})) - \varphi(\widehat{v})|\nabla f(\overline{x})\cdot\nabla_{\overline{y}}(\xi\,\lambda)(t,\overline{x},f(\overline{x}),y)dy\,d\overline{x}\,dt \\
&\quad -\int_{[0,T)\times\Pi}\int_\Omega |\varphi(\overline{u}_\Sigma(t,\overline{x})) - \varphi(\widehat{v})|\partial_{y_d}\,(\xi\,\lambda)(t,\overline{x},f(\overline{x}),y)\,dy\,d\overline{x}\,dt\,dy \\
&= -\widetilde{J^x_{\rho_m}} + \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\overline{u}_\Sigma) - \varphi(\widehat{v})|\rho_m(\overline{x}-\overline{y})\,\rho_n'(f(\overline{x})-y_d)\,(\psi\,\lambda)_{\Sigma_x}dy\,d\overline{x}\,dt
\end{aligned}$$

so that

$$\begin{aligned}
\widetilde{J^x} + \widetilde{J^y} &= \widetilde{J^x_f} + \widetilde{J^x_\psi} \\
&\quad + \int_{[0,T)\times\Pi}\int_\Omega |\varphi(\overline{u}_\Sigma) - \varphi(\widehat{v})|\,(1+|\nabla f(\overline{x})|^2)\,\rho_m(\overline{x}-\overline{y})\rho_n'(f(\overline{x})-y_d)\,(\psi\,\lambda)_{\Sigma_x}\,d\overline{x}\,dt\,dy.
\end{aligned}$$

In particular, no derivatives of ρ_m appear in $J^x + J^y$. Hence, summing up by $\widetilde{v}$ the quantity $v(t,\overline{x},y_d,\beta)$ and passing to the limit $[m\to+\infty]$ in $\lim_{\varepsilon\to 0}(A+B) = \widetilde{I} + \widetilde{J^x} + \widetilde{J^y}$, we get:

$$\lim_{m\to+\infty}\ \lim_{\varepsilon\to 0}(A+B) = \overline{I} + \overline{J_f} + \overline{J_\psi} + \overline{J_{\rho_n}}\,,$$

with

$$\overline{I} = -\int_{[0,T)\times\Pi\times(0,1)}\int_0^\infty\int_0^1 (\Phi(t,\overline{x},f(\overline{x}),\widetilde{v},\overline{u}_\Sigma)\cdot\begin{pmatrix} -\nabla f(\overline{x}) \\ 1 \end{pmatrix}\,\rho_n(f(\overline{x})-y_d)\,(\psi\,\lambda)_{\Sigma_x}\,d\overline{x}\,dt\,d\alpha\,dy_d\,d\beta\,,$$

$$\overline{J_f} = \int_{[0,T)\times\Pi}\int_0^\infty \quad |\varphi(\widetilde{v}) - \varphi(\overline{u}_\Sigma)|\Delta f(\overline{x})\,(\psi\,\lambda)_{\Sigma_x}\,\rho_n(f(\overline{x})-y_d)\,d\overline{x}\,dt\,dy_d\,,$$

$$\overline{J_\psi} = \int_{[0,T)\times\Pi}\int_0^\infty \quad |\varphi(\widetilde{v}) - \varphi(\overline{u}_\Sigma)|\nabla f(\overline{x})\cdot\nabla_{\overline{x}}\left((\psi\,\lambda)_{\Sigma_x}\right)\rho_n(f(\overline{x})-y_d)\,d\overline{x}\,dt\,dy_d\,,$$

$$\overline{J_{\rho_n}} = \int_{[0,T)\times\Pi}\int_0^\infty \quad |\varphi(\widetilde{v}) - \varphi(\overline{u}_\Sigma)|\,(1+|\nabla f(\overline{x})|^2)\,\rho_n'(f(\overline{x})-y_d)\,(\psi\,\lambda)_{\Sigma_x}\,d\overline{x}\,dt\,dy_d\,.$$

To compute the limit as n tends to $+\infty$ of the four preceding terms, first recall that $\mathrm{trace}((\varphi(v)) - \varphi(\overline{u}_\Sigma)) = 0$, and therefore:

$$\lim_{n\to+\infty}\overline{J_f} = 0 \ \text{ and } \ \lim_{n\to+\infty}\overline{J_\psi} = 0\,.$$

Besides, remark that

$$\Delta\,\omega_{1/n}(x) = -\rho_n'(f(\overline{x}) - x_d)\,(1 + \nabla f(\overline{x})|^2) + \rho_n(f(\overline{x}) - x_d)\,\Delta f(\overline{x})\,,$$

so that, replacing y_d by x_d in $\overline{J_{\rho_n}}$, we have:

$$\begin{aligned}\overline{J_{\rho_n}} &= -\int_Q |\varphi(v) - \varphi(\overline{u}_\Sigma(t,\overline{x}))|\,\Delta\,\omega_{1/n}(x)\,(\psi\,\lambda)(t,\overline{x},f(\overline{x}))\,dx\,dt + \overline{J_f}\\ &= \int_Q \nabla|\varphi(v) - \varphi(\overline{u}_\Sigma(t,\overline{x}))|\,\nabla\,\omega_{1/n}(x)\,(\psi\,\lambda)(t,\overline{x},f(\overline{x}))\,dx\,dt + \overline{\varepsilon_n^1}\,.\end{aligned}$$

Here, $\overline{\varepsilon_n^1} = \overline{J_f} + \displaystyle\int_Q |\varphi(v) - \varphi(\overline{u}_\Sigma(t,\overline{x}))|\,\nabla\omega_{1/n}(x)\cdot\nabla(\psi\,\lambda)_{\Sigma_x}\,dx\,dt$ tends to zero when $n \to +\infty$. Moreover

$$\overline{I} = -\int_Q \Phi(t,x,v,\overline{u}_\Sigma)\cdot\nabla\omega_{1/n}(x)\,(\psi\,\lambda)_{\Sigma_x}\,dx\,dt\; + \overline{\varepsilon_n^2}\,,$$

where $\overline{\varepsilon_n^2} = \displaystyle\int_Q (\Phi(t,x,v,\overline{u}_\Sigma) - \Phi(t,\overline{x},f(\overline{x}),v,\overline{u}_\Sigma)\cdot\nabla\omega_{1/n}(x)\,(\psi\,\lambda)_{\Sigma_x}\,dx\,dt$ tends to zero when $n \to +\infty$.
Using Formula (3.10) we get

$$\liminf_{n\to+\infty}\lim_{m\to+\infty}\lim_{\varepsilon\to 0}(A+B) = -\limsup_{n\to+\infty}\int_Q \mathcal{G}_x(t,x,v(t,x,\beta),\overline{u}_\Sigma)\cdot\nabla\omega_{1/n}(x)\,(\psi\,\lambda)_\Sigma\,dx\,dt\,d\beta\,.$$

Starting from Inequality (3.21) and taking the limit with respect to m, then the infimum limit with respect to n of both sides yields:

$$\int_{Q^1}\int_0^1 \left[|u-v|\,(\psi\lambda)_t + \mathcal{G}_x(t,x,u,v)\cdot\nabla(\psi\lambda)\right]\,dx\,dt\,d\alpha\,d\beta \geq$$

$$\left[\begin{array}{l} -\displaystyle\limsup_{n\to+\infty}\int_{Q^1} \mathcal{G}_x(t,x,v(t,x,\beta),\overline{u}_\Sigma(t,\overline{x}))\cdot\nabla\omega_{1/n}\,(\psi\,\lambda)(t,\overline{x},f(\overline{x}))\,dx\,dt\,d\beta\\ +\displaystyle\limsup_{\varepsilon\to 0}\int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma(t,\overline{x}))\cdot\nabla\omega_\varepsilon(x)\,(\psi\,\lambda)(t,x)\,dx\,dt\,d\alpha\\ +\displaystyle\lim_{n\to+\infty}\lim_{m\to+\infty} E\,. \end{array}\right]$$

Clearly, $\displaystyle\lim_{n\to+\infty}\lim_{m\to+\infty} E = 0$ and by Remark 3.4.1 we can replace $(\psi\lambda)_\Sigma$ by $(\psi\lambda)$ in the previous $\limsup$ in n , so if we denote by $A(u,v)$ the quantity

$$\begin{aligned}A(u,v) &= -\limsup_{\varepsilon\to 0}\int_{Q^1} \mathcal{G}_x(t,x,v(t,x,\beta),\overline{u}_\Sigma)\cdot\nabla\omega_\varepsilon(x)\,\psi\,\lambda\,dx\,dt\,d\beta\\ &\qquad +\limsup_{\varepsilon\to 0}\int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma)\cdot\nabla\omega_\varepsilon(x)\,\psi\,\lambda\,dx\,dt\,d\alpha\,,\end{aligned}$$

we get

$$\int_{Q^1}\int_0^1 \left[|u-v|\,(\psi\lambda)_t + \mathcal{G}_x(t,x,u,v)\cdot\nabla(\psi\lambda)\right]\,dx\,dt\,d\alpha\,d\beta \geq A(u,v)\,. \tag{3.22}$$

Conclusion

Suppose the following lemma already proved

Lemma 3.4.2 *Let $u \in L^\infty(Q^1)$ be an entropy process solution of Problem (3.1). Then, for any non-negative function $\psi \in \mathcal{C}_c^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$,*

$$\begin{aligned}-\infty \quad &< \quad \liminf_{\varepsilon\to 0} \int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma) \cdot \nabla\omega_\varepsilon(x)\, \psi\, \lambda \, dx\, dt\, d\alpha \\ &\leq \quad \limsup_{\varepsilon\to 0} \int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma) \cdot \nabla\omega_\varepsilon(x)\, \psi\, \lambda \, dx\, dt\, d\alpha \ \leq 0\,.\end{aligned}$$

Remark 3.4.2 *It is in the course of the proof of the estimate*

$$-\infty < \liminf_{\varepsilon\to 0} \int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}_\Sigma) \cdot \nabla\omega_\varepsilon(x)\, \psi\, \lambda \, dx\, dt\, d\alpha$$

(and not elsewhere) that assumption $\Delta\varphi(\overline{u}_\Sigma) \in L^1(Q)$ is required. We do not know how to remove, or weaken, this hypothesis.

As a consequence of Lemma 3.4.2, the function A is antisymmetric in (u,v), while the left hand-side of Inequality (3.22) is a symmetric function of (u,v). Thus, the following inequality holds:

$$\int_{Q^1}\int_0^1 \left[|u-v|\,(\psi\lambda)_t + \mathcal{G}_x(t,x,u,v)\cdot\nabla(\psi\lambda)\right]\, dx\, dt\, d\alpha\, d\beta \geq 0\,.$$

Now, recall that $\lambda = \lambda_\alpha$ is an element of the partition of unity $(\lambda_\alpha)_{0\leq\alpha\leq N}$: summing the previous inequality over $\alpha \in 0,\ldots,N$ yields

$$\int_{Q^1}\int_0^1 \left[|u-v|\,\psi_t + \mathcal{G}_x(t,x,u,v)\cdot\nabla\psi\right]\, dx\, dt\, d\alpha\, d\beta \geq 0\,. \tag{3.23}$$

Define a positive function ψ_0 by

$$\psi_0(t,x) = \psi_0(t) = (T-t)\chi_{(0,T)}(t)\,.$$

Applying (3.23) with ψ_0 as a test function yields:

$$\int_0^T\int_\Omega\int_0^1\int_0^1 |u(t,x,\alpha) - v(t,x,\beta)|\, dx\, dt\, d\alpha\, d\beta \ \leq 0\,.$$

Consequently, there holds $u(t,x,\alpha) = v(t,x,\beta)$ for a.e. $(t,x,\alpha,\beta) \in Q\times(0,1)\times(0,1)$. Defining the function w by the formula

$$w(t,x) = \int_0^1 u(t,x,\alpha)\, d\alpha\,,$$

and accounting for the product structure of the measurable space $Q \times (0,1) \times (0,1)$ we get:

$$u(t,x,\alpha) = w(t,x) = v(t,x,\beta) \ \textit{for a.e.}\ (t,x,\alpha,\beta) \in Q \times (0,1)^2.$$

There remains to prove Lemma 3.4.2: set

$$g = \partial_t \overline{u}_\Sigma + \mathrm{div}_x F(t,x,\overline{u}_\Sigma) - \Delta\varphi(\overline{u}_\Sigma).$$

From the hypothesis made on $\overline{u}$, follows: $g \in L^1([0,T) \times B \cap Q)$ and the function $\overline{u}$ can be seen as an entropy solution of the equation $\partial_t w + \mathrm{div}_x F(t,x,w) - \Delta\varphi(w) = g$ with unknown w with a certain boundary condition that is not important. Now, we use a result of comparison and assert that, for any non-negative function $\theta \in \mathcal{C}_c^\infty(\mathrm{I\!R}_+ \times B \cap Q)$, there holds

$$\int_{Q^1} [|u-\overline{u}|\,\theta_t + \mathcal{G}_x(t,x,u,\overline{u})) \cdot \nabla\theta]\, dx\,dt\,d\alpha + \int_\Omega |u_0 - \overline{u}(0,x)|\,\theta(0)\,dx \\ + \int_{Q^1} \mathrm{sgn}(u-\overline{u})\, g\,\theta\, dx\,dt\,d\alpha \geq 0.$$

This result of comparison can easily be deduced from the preceding work. Indeed, no care about the boundary condition is needed, for the function $\theta(t,\cdot)$ has compact support included in Ω. Now applying this result with $\theta(x,t) = \omega_\varepsilon(x)\lambda(x)\psi(t,x)$ yields

$$\int_{Q^1} \mathcal{G}_x(t,x,u,\overline{u}) \cdot \nabla\omega_\varepsilon\, \psi\lambda\, dx\,dt\,d\alpha \geq - \left[\begin{array}{l} \displaystyle\int_{Q^1} [|u-\overline{u}|\,\psi_t\,\omega_\varepsilon\,\lambda + \mathcal{G}_x(t,x,u,\overline{u}) \cdot \nabla(\psi\lambda)\,\omega_\varepsilon]\, dx\,dt\,d\alpha \\ \displaystyle\int_\Omega |u_0 - \overline{u}(0,x)|\,\psi(0)\,\lambda\,\omega_\varepsilon\,dx \\ \displaystyle + \int_{Q^1} \mathrm{sgn}(u-\overline{u})\, g\,\psi\,\lambda\,\omega_\varepsilon\,, dx\,dt\,d\alpha \end{array} \right].$$

As the function ω_ε is bounded, the right hand-side of this inequality is bounded from below, independently of ε. Lemma 3.4.2 follows.

3.5 The finite volume scheme

The mesh used to discretize Problem (3.1) has to be structured enough in order to ensure the consistency of the fluxes, mainly because a second order problem is consider (at least when the function φ is not constant). This is specified in the following section.

3.5.1 Assumptions and notations

Definition 3.5.1 (Admissible mesh of Ω) *An admissible mesh of Ω is given by a set $\mathcal{T}$ of open bounded polygonal convex subsets of Ω called control volumes, a family $\mathcal{E}$ of subsets of $\bar{\Omega}$ contained in hyper planes* *of $\mathbb{R}^d$ with strictly positive measure, and a family of points (the "centers" of control volumes) satisfying the following properties:*

(i) The closure of the union of all control volumes is $\bar{\Omega}$.

(ii) For any $K \in \mathcal{T}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \bar{K} \backslash K = \cup_{\sigma \in \mathcal{E}_K} \bar{\sigma}$. Furthermore, $\mathcal{E} = \cup_{K \in \mathcal{T}} \mathcal{E}_K$.

(iii) For any $(K, L) \in \mathcal{T}^2$ with $K \neq L$, either the "length" (i.e. the $(d-1)$-dimensional Lebesgue measure) of $\bar{K} \cap \bar{L}$ is 0 or $\bar{K} \cap \bar{L} = \bar{\sigma}$ for some $\sigma \in \mathcal{E}$. In the latter case, we shall write $\sigma = K|L$ and $\mathcal{E}_{int} = \{\sigma \in \mathcal{E}, \exists (K, L) \in \mathcal{T}^2, \sigma = K|L\}$. For any $K \in \mathcal{T}$, we shall denote by $\mathcal{N}_K$ the set of boundary control volumes of K, i.e. $\mathcal{N}_K = \{L \in \mathcal{T}, K|L \in \mathcal{E}_K\}$.

(iv) The family of points $(x_K)_{K \in \mathcal{T}}$ is such that $x_K \in K$ (for all $K \in \mathcal{T}$) and, if $\sigma = K|L$, it is assumed that the straight line (x_K, x_L) is orthogonal to σ.
For a control volume $K \in \mathcal{T}$, we will denote by $m(K)$ its measure and by $\mathcal{E}_{ext,K}$ the subset of the edges of K included in the boundary $\partial\Omega$. If $L \in \mathcal{N}_K$, $m(K|L)$ will denote the measure of the edge between K and L, $\tau_{K|L}$ the "transmissibility" through $K|L$, defined by $\tau_{K|L} = \dfrac{m(K|L)}{d(x_k, x_L)}$. Similarly, if $\sigma \in \mathcal{E}_{ext,K}$, we will denote by $m(\sigma)$ its measure and τ_σ the "transmissibility" through σ, defined by $\tau_\sigma = \dfrac{m(\sigma)}{d(x_K, \sigma)}$. One denotes $\mathcal{E}_{ext} = \cup_{K \in \mathcal{T}} \mathcal{E}_{ext,K}$. The size of the mesh $\mathcal{T}$ is defined by

$$\mathrm{size}(\mathcal{T}) = \max_{K \in \mathcal{T}} \mathrm{diam}(K),$$

and a geometrical factor, linked with the regularity of the mesh, is defined by

$$\mathrm{reg}(\mathcal{T}) = \min_{K \in \mathcal{T}, \sigma \in \mathcal{E}_K} \frac{d(x_K, \sigma)}{\mathrm{diam}(K)}.$$

Remark 3.5.1 *Examples of meshes satisfying these assumptions are triangular meshes satisfying the acute angle condition (in fact this condition may be weakened to the Delaunay condition), rectangular meshes or Voronoï meshes, see [EGH99] or [EGH00b] for more details.*

Definition 3.5.2 (Time discretization of $(0, T)$) *A time discretization of $(0, T)$ is given by an integer value N and by an increasing sequence of real values $(t^n)_{n \in [\![0, N+1]\!]}$ with $t^0 = 0$ and $t^{N+1} = T$. The time steps are then defined by $\delta t^n = t^{n+1} - t^n$, for $n \in [\![0, N]\!]$.*

Definition 3.5.3 (Space-time discretization of Q) *A finite volume discretization D of Q is the family $D = (\mathcal{T}, \mathcal{E}, (x_K)_{K \in \mathcal{T}}, N, (t^n)_{n \in [\![0, N]\!]})$, where $\mathcal{T}$, $\mathcal{E}$, $(x_K)_{K \in \mathcal{T}}$ is an admissible mesh of Ω in the sense of Definition 3.5.1 and N, $(t^n)_{n \in [\![0, N+1]\!]}$ is a time discretization of $(0, T)$ in the sense of Definition 3.5.2. For a given finite volume discretization D, one defines:*

$$\mathrm{size}(D) = \max(\mathrm{size}(\mathcal{T}), (\delta t^n)_{n \in [\![0, N]\!]}), \qquad \textit{and } \mathrm{reg}(D) = \mathrm{reg}(\mathcal{T}).$$

3.5.2 The scheme

We may now define the finite volume discretization of (3.1). Let D be a finite volume discretization of Q in the sense of Definition 3.5.3. The initial condition is discretized by

$$U_K^0 = \frac{1}{m(K)} \int_K u_0(x)dx, \quad \forall K \in \mathcal{T}. \tag{3.24}$$

In order to introduce the finite volume scheme, we need to define:

$$\bar{U}_\sigma^{n+1} = \frac{1}{\delta t^n \; m(\sigma)} \int_{t^n}^{t^{n+1}} \int_\sigma \bar{u}(t,x)d\gamma(x)dt, \qquad \forall \sigma \in \mathcal{E}_{ext}, \forall n \in [\![0,N]\!]. \tag{3.25}$$

An **implicit finite volume scheme** for the discretization of Problem (3.1) is given by the following set of non-linear equations with unknowns $U_D = (U_K^n)_{K\in\mathcal{T},n\in[\![0,N+1]\!]}$:

$\forall K \in \mathcal{T}, \forall n \in [\![0,N]\!]$,

$$\frac{U_K^{n+1} - U_K^n}{\delta t^n}\; m(K) + \sum_{\sigma\in\mathcal{E}_K} m(\sigma) F_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma(\varphi(U_{K,\sigma}^{n+1}) - \varphi(U_K^{n+1})) = 0 \tag{3.26}$$

where

$$U_{K,\sigma}^{n+1} = \begin{cases} U_L^{n+1} if\ \sigma = K|L \\ \bar{U}_\sigma^{n+1} if\ \sigma \in \mathcal{E}_{ext} \end{cases} \tag{3.27}$$

and $F_{K,\sigma}^{n+1}$ is a monotonous flux consistent with F which means that:

- $\forall v \in \mathbb{R}$, $u \mapsto F_{K,\sigma}^{n+1}(u,v)$ is nondecreasing and $\forall u \in \mathbb{R}$, $v \to F_{K,\sigma}^{n+1}(u,v)$ is non increasing
- $F_{K,\sigma}^{n+1}(u,v) = -F_{K,\sigma}^{n+1}(v,u)$ for all $(u,v) \in \mathbb{R}^2$
- $F_{K,\sigma}^{n+1}$ is a locally Lipschitz continuous function
- $F_{K,\sigma}^{n+1}(s,s) = \frac{1}{\delta t_n}\frac{1}{\sigma}\int_{t^n}^{t^{n+1}}\int_\sigma F(x,t,s)\cdot \mathbf{n}_{K,\sigma} d\gamma(x)\delta t.$

The Godunov scheme and the splitting flux scheme of Osher may be the most popular examples of monotone fluxes.

3.6 Monotonicity of the scheme and direct consequences

As we said in the introduction, one of the difficulties of this problem is the definition of a physically admissible solution. The monotonous schemes are well known to add numerical viscosity to the equations. They are L^∞ stable and they respect discrete entropy inequalities. In other words, continuous entropy inequality have their discrete analogue and they are respected by any solution of (3.24)-(3.27). This is summarized in the following proposition:

Proposition 3.6.1 (Discrete entropy inequalities and consequences) *There exists a unique solution to the scheme. Moreover, it satisfies the following maximum principle property and entropy inequalities:*
$\forall K \in \mathcal{T}, \forall n \in [\![0, N]\!]$,

$$A \leq U_K^{n+1} \leq B, \tag{3.28}$$

$$\begin{aligned} \frac{\eta_\kappa^+(U_K^{n+1}) - \eta_\kappa^+(U_K^n)}{\delta t^n} m(K) \quad &+ \quad \sum_{\sigma \in \mathcal{E}_K} m(\sigma) \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) \\ &- \quad \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma (\eta_{\varphi(\kappa)}^+(\varphi(U_{K,\sigma}^{n+1})) - \eta_{\varphi(\kappa)}^+(\varphi(U_K^{n+1}))) \leq 0, \end{aligned} \tag{3.29}$$

and

$$\begin{aligned} \frac{\eta_\kappa^-(U_K^{n+1}) - \eta_\kappa^-(U_K^n)}{\delta t^n} m(K) \quad &+ \quad \sum_{\sigma \in \mathcal{E}_K} m(\sigma) \Phi_{K,\sigma,\kappa}^{-,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) \\ &- \quad \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma (\eta_{\varphi(\kappa)}^-(\varphi(U_{K,\sigma}^{n+1})) - \eta_{\varphi(\kappa)}^-(\varphi(U_K^{n+1}))) \leq 0. \end{aligned} \tag{3.30}$$

where $\Phi_{K,\sigma,\kappa}^{+,n+1}$ and $\Phi_{K,\sigma,\kappa}^{-,n+1}$ are discrete fluxes defined by

$$\begin{aligned} \Phi_{K,\sigma,\kappa}^{+,n+1}(u,v) &= F_{K,\sigma}^{n+1}(u \top \kappa, v \top \kappa) - F_{K,\sigma}^{n+1}(\kappa, \kappa), \\ \Phi_{K,\sigma,\kappa}^{-,n+1}(u,v) &= F_{K,\sigma}^{n+1}(\kappa, \kappa) - F_{K,\sigma}^{n+1}(u \bot \kappa, v \bot \kappa). \end{aligned}$$

Proof. Since it is an implicit problem, we will first show that if U_D is a solution to the scheme, it satisfies the discrete inequalities (3.29) and (3.30). Then we obtain an a priori maximum principle that allows us to obtain existence by Leray-Schauder theorem. Then we prove uniqueness by an analogue method as we use to prove discrete entropy inequalities.

For $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$, we denote by $B_K^{n+1}(U_K^{n+1}, U_K^n, (U_L^{n+1})_{L \in \mathcal{N}(K)}, (\bar{U}_\sigma^{n+1})_{\sigma \in \mathcal{E}_{ext,K}})$ the first member in Equation (3.26). The monotony of the scheme imply that B_K^{n+1} is non increasing with respect to all its arguments except the first one U_K^{n+1} and for every $\kappa \in \mathbb{R}$ we obtain:

$$B_K^{n+1}(\kappa, \kappa, (\kappa)_{L \in \mathcal{N}(K)}, (\kappa)_{\sigma \in \mathcal{E}_{ext,K}}) = 0 \tag{3.31}$$

So

$$\begin{aligned}B_K^{n+1}(U_K^{n+1}, U_K^n \top \kappa, (U_L^{n+1} \top \kappa)_{L\in\mathcal{N}(K)}, (\bar{U}_\sigma^{n+1} \top \kappa)_{\sigma\in\mathcal{E}_{ext,K}}) &\leq B(U_K^{n+1}, U_K^n, (U_L^{n+1})_{L\in\mathcal{N}(K)}, (\bar{U}_\sigma^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}) \\ &= 0\end{aligned}$$

and

$$\begin{aligned}B_K^{n+1}(\kappa, U_K^n \top \kappa, (U_L^{n+1} \top \kappa)_{L\in\mathcal{N}(K)}, (\bar{U}_\sigma^{n+1} \top \kappa)_{\sigma\in\mathcal{E}_{ext,K}} &\leq B(\kappa, \kappa, (\kappa)_{L\in\mathcal{N}(K)}, (\kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \\ &= 0\end{aligned}$$

Now since $(U_K^{n+1} \top \kappa)$ equals U_K^{n+1} or κ, we have

$$B_K^{n+1}(U_K^{n+1} \top \kappa, U_K^n \top \kappa, (U_L^{n+1} \top \kappa)_{L\in\mathcal{N}(K)}, (\bar{U}_\sigma^{n+1} \top \kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \leq 0. \tag{3.32}$$

By an analog proof we also have

$$B_K^{n+1}(U_K^{n+1} \perp \kappa, U_K^n \perp \kappa, (U_L^{n+1} \perp \kappa)_{L\in\mathcal{N}(K)}, (\bar{U}_\sigma^{n+1} \perp \kappa)_{\sigma\in\mathcal{E}_{ext,K}}) \geq 0. \tag{3.33}$$

Then by using (3.3) and (3.4), Inequalities (3.29) and (3.30) are direct consequences of (3.31),(3.32) and (3.31),(3.33) respectively.

Now, if we take $\kappa = \max(U_K^n, (U_{K,\sigma}^{n+1})_\sigma \in \mathcal{E}_K)$, we easily get from discrete entropy inequalities (3.29) that $U_K^{n+1} \leq \kappa$. So by induction, we obtain

$$U_K^{n+1} \leq \max(ess \sup_\Omega(u_0), \max_{\partial\Omega}(\bar{u})) \leq B.$$

By the same argument, we have also $U_K^{n+1} \geq A$ and we proved maximum principle (3.28).

Now it is not very difficult to obtain existence of a solution, using Leray-Schauder theorem. Indeed, let us take $n \in [\![0, N]\!]$, replace $U_{K,\sigma}^{n+1}$ by its value and assume that U_K^n is already known and belongs to the interval $[A, B]$. Then the problem is to find $U_\mathcal{T}^{n+1} = (U_K^{n+1})_{K\in\mathcal{T}}$ satisfying the equation $f(U_\mathcal{T}^{n+1}) = 0$, where $f : (U_K^{n+1})_{K\in\mathcal{T}} \mapsto (B_K^{n+1})_{K\in\mathcal{T}}$.

Let us define a continuous deformation f^t of f by replacing $F_{K,\sigma}^{n+1}$ by $tF_{K,\sigma}^{n+1}$ and φ by $t\varphi$. Then, $(t, U_\mathcal{T}^{n+1}) \mapsto f^t(U_\mathcal{T}^{n+1})$ is continuous. The scheme defined by f^t is also monotone , then we also have the maximum principle and for $\varepsilon > 0$, the equation $f^t(U_\mathcal{T}^{n+1}) = 0$ has no solution in the boundary of the open subset $X = (A-\varepsilon, B+\varepsilon)^\mathcal{T}$ of $\mathbb{R}^\mathcal{T}$. Moreover, f^0 is a linear function so $degree(f^0, X) \neq 0$. Then by Leray-Schauder theorem,

$$degree(f, X) = degree(f^0, x) \neq 0.$$

Hence the problem $f(U_\mathcal{T}^{n+1}) = 0$ has at least one solution and, by induction, there exist a solution to the scheme (3.24)-(3.27).

Let us now prove uniqueness. Let U_D and V_D be two solutions of the scheme . By the same arguments involved in the proof of (3.29) and (3.30), we have

$$B(U_K^{n+1}\top V_K^{n+1}, U_K^n \top V_K^n, (U_L^{n+1}\top V_L^{n+1})_{L\in\mathcal{N}(K)}, (\bar{U}_\sigma^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}) \leq 0$$

and

$$B(U_K^{n+1}\bot V_K^{n+1}, U_K^n \bot V_K^n, (U_L^{n+1}\bot V_L^{n+1})_{L\in\mathcal{T}}, (\bar{U}_\sigma^{n+1})_{\sigma\in\mathcal{E}_{ext,K}}) \geq 0.$$

So we get $\forall K \in \mathcal{T}, n \in [\![0, N]\!]$,

$$\begin{aligned}\frac{|U_K^{n+1} - V_K^{n+1}| - |U_K^n - V_K^n|}{\delta t_n} m(K) \quad &+ \quad \sum_{\sigma\in\mathcal{E}_K} H_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}, V_K^{n+1}, V_{K,\sigma}^{n+1}) \\ &- \quad \sum_{\sigma} \tau_\sigma(|\varphi(U_K^{n+1}) - \varphi(V_K^{n+1})| - |\varphi(U_{K,\sigma}^{n+1} - \varphi(V_{K,\sigma}^{n+1})|) \leq 0\end{aligned}$$

where

$$H_{K,\sigma}^{n+1}(u, v, \kappa, \rho) = F_{K,\sigma}^{n+1}(u\top\kappa) - F_{K,\sigma}^{n+1}(v\bot\rho)$$

Now if we had the following inequalities on $K \in \mathcal{T}$, by using that the scheme is conservative, we obtain $\forall n \in [\![0, N]\!]$,

$$\sum_{K\in\mathcal{T}} \frac{|U_K^{n+1} - V_K^{n+1}| - |U_K^n - V_K^n|}{\delta t_n} m(K) \leq 0,$$

Then since $U_K^0 - V_K^0 = 0$, we have $U_K^{n+1} - V_K^{n+1} = 0$ for every $n \in \mathcal{T}$ and the proof is complete.

3.7 A priori estimates

The inequalities we derive from monotony and local conservation are L^∞ and L^1 estimates. We will prove now L^2 estimates. We introduce a discretization $\bar{U}_D = (\bar{U}_K^{n+1})_{\{K\in\mathcal{T}, n\in[\![0,N]\!]\}}$ of $\bar{u}$:

$$\bar{U}_K^{n+1} = \frac{1}{\delta t_n}\frac{1}{m(K)}\int_{t^n}^{t^{n+1}}\int_K \bar{u}, \qquad \forall K \in \mathcal{T}, \forall n \in [\![0, N]\!]$$

Proposition 3.7.1 **($L^2(0, T, H^1(\Omega))$ and weak BV estimate)** *Let U_D be the solution of (3.24)-(3.27)(cf Proposition 3.6.1), assume that $reg(D) \leq \xi$ then there exist $C(\xi, T, \Omega, \Phi, \max(Lip(F_{K,\sigma}^{n+1})), \bar{u}, A, B)$ such that*

$$(\mathcal{N}_D(\zeta(u_D)))^2 = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} (\frac{1}{2}\sum_{\sigma\in\mathcal{E}_{int,K}} \tau_\sigma(\zeta(U_K^{n+1}) - \zeta(U_{K,\sigma}^{n+1}))^2 + \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma(\zeta(U_K^{n+1}) - \zeta(U_{K,\sigma}^{n+1}))^2) \leq C$$

and

$$(wBVF_D(u_D))^2 = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} m(\sigma) \max_{U_K^{n+1}\le c\le d\le U_{K,\sigma}^{n+1}} ((F(d,c)-F(d,d))^2 + (F(d,c)-F(c,c))^2) \le C.$$

Remark 3.7.1 *The notation $wBVF_D$ is for discrete "weak BV inequality" on $F(t,x,u)$. See [EGH00a], [EGH00b] or [CH99].*

Let us multiply the equations by $\delta t_n(U_K^{n+1} - \bar{U}_K^{n+1})$ and sum over $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$. Then it holds

$$E1 + E2 + E3 = 0,$$

with

$$E1 = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} (U_K^{n+1} - U_K^n)(U_K^{n+1} - \bar{U}_K^{n+1})m(K)$$

$$E2 = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1})(U_K^{n+1} - \bar{U}_K^{n+1})$$

and

$$E3 = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma(\varphi(U_K^{n+1}) - \varphi(U_{K,\sigma}^{n+1}))(U_K^{n+1} - \bar{U}_K^{n+1})$$

By a discrete time integrate by parts, we get

$$E1 \ge \sum_{K\in\mathcal{T}} m(K)\frac{1}{2}(U_K^{N+1} - \bar{U}_K^{N+1})^2 - \sum_{K\in\mathcal{T}} m(K)\frac{1}{2}(U_K^0 - \bar{U}_K^0)^2 - E1,1$$

where

$$\begin{aligned} E1,1 &= \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{|\bar{U}_K^{n+1} - \bar{U}_K^n|}{\delta t_n} |U_K^{n+1} - \bar{U}_K^{n+1}| \\ &\le (B-A)\|\bar{u}_t\|_{L^1(Q)} \end{aligned}$$

By a discrete space integrate by parts, using the fact that $U_D - \bar{U}_D$ is equal to zero on the border, we obtain

$$E3 = E3int + E3ext + E3,1$$

where

$$\begin{aligned}
E3int &= \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} \tau_\sigma (U_K^{n+1} - U_{K,\sigma}^{n+1})(\varphi(U_K^{n+1}) - \varphi(U_{K,\sigma}^{n+1})),\\
E3ext &= \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma (U_K^{n+1} - U_{K,\sigma}^{n+1})(\varphi(U_K^{n+1}) - \varphi(U_{K,\sigma}^{n+1})),\\
E3,1 &= \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} (\frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} \tau_\sigma (U_K^{n+1} - U_{K,\sigma}^{n+1})(\varphi(U_K^{n+1}) - \varphi(U_{K,\sigma}^{n+1})).
\end{aligned}$$

If we use the property $(\zeta(a) - \zeta(b))^2 \leq (a-b)(\varphi(a) - \varphi(b))$ we get

$$\begin{aligned}
E3int &\geq \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} \tau_\sigma (\zeta(U_K^{n+1}) - \zeta(U_{K,\sigma}^{n+1}))\\
E3ext &\geq \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma (\zeta(U_K^{n+1}) - \zeta(U_{K,\sigma}^{n+1}))
\end{aligned}$$

For the last term $E3,1$, by inequality $\varphi(a) - \varphi(b) \leq \sqrt{\Phi}(\zeta(a) - \zeta(b))$, we obtain

$$E3,1 \quad \geq \quad -\sqrt{\Phi}\mathcal{N}_D(\bar{U}_D)\mathcal{N}_D(\zeta(U_D)),$$

To get a bound on $\mathcal{N}_D(\bar{U}_D)$, we use the following inequality proved in [EGH99]:

$$\mathcal{N}_D(\bar{u}) \leq C(reg(D))\|\nabla \bar{u}\|_{L^2(Q)}.$$

Now let us deal with $E2$. By using the assumption $\text{div}_x(F(x,t,u)) = 0$, we obtain

$$\sum_{\sigma\in\mathcal{E}_K} m(\sigma) F_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) = \sum_{\sigma\in\mathcal{E}_K} m(\sigma)(F_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - F_{K,\sigma}^{n+1}(U_K^{n+1}, U_K^{n+1}))$$

By gathering by edges, if we denote by $E2,1$ the first term in $E2$ (the term which only concern U_D) we get

$$\begin{aligned}
E2,1 = &\sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} m(\sigma)(U_K^{n+1}(F_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - F_{K,\sigma}^{n+1}(U_K^{n+1}, U_K^{n+1}))\\
&-U_{K,\sigma}^{n+1}(F_{K,\sigma}^{n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - F_{K,\sigma}^{n+1}(U_{K,\sigma}^{n+1}, U_{K,\sigma}^{n+1}))) + E2,1ext,
\end{aligned}$$

where

$$\begin{aligned} E2,1ext &= \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} m(\sigma)(F_{K,\sigma}^{n+1}(U_K^{n+1},U_{K,\sigma}^{n+1}) - F_{K,\sigma}^{n+1}(U_K^{n+1},U_K^{n+1}))U_K^{n+1} \\ &\geq -\max(|A|,|B|)Lip(F_{K,\sigma}^{n+1})(B-A)\,T\,m(\partial\Omega). \end{aligned}$$

Then for $E2,1$, it remains to deal with $E2,1int = E2,1 - E2,1ext$. For that, let us apply the method of Claire Chainais-Hillairet [CH99]. We introduce the function $G_{K,\sigma}^{n+1}$ as a primitive of $s\frac{d}{ds}F_{K,\sigma}^{n+1}(s,s)$. It is not difficult to prove the following lemma:

Lemma 3.7.1 (Technical lemma) *If G is a primitive of $s\frac{d}{ds}F(s,s)$ and F is a Lipschitz function, non-decreasing with respect to its first argument and non increasing with respect to the other argument then*

$$\begin{aligned} G(b)-G(a)-b(F(b,b)-F(a,b))+a(F(a,a)-F(a,b)) \quad \geq \quad & \frac{1}{4\ Lip(F)}(\max_{a\leq c\leq d\leq b}(F(d,c)-F(d,d))^2 \\ & + \max_{a\leq c\leq d\leq b}(F(d,c)-F(c,c))^2) \end{aligned}$$

By using this lemma and the fact that $\displaystyle\sum_{\sigma\in\mathcal{E}_K} G_{K,\sigma}^{n+1}(s)=0$, and $G_{K,\sigma}^{n+1}(s) = -G_{L,\sigma}^{n+1}(s)$ if $\sigma = K|L$, we get

$$E2,1int \geq \frac{1}{4Lip(F_{K,\sigma}^{n+1})}(wBVF_D(U_D))^2$$

For $E2,2$, we perform an integrate by parts

$$E2,2 = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} m(\sigma)(F_{K,\sigma}^{n+1}(U_K^{n+1},U_{K,\sigma}^{n+1}) - F_{K,\sigma}^{n+1}(U_K^{n+1},U_K^{n+1}))(\bar{U}_K^{n+1}-\bar{U}_\sigma^{n+1}) + E2,2ext$$

where

$$\begin{aligned} E2,2ext &= \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} m(\sigma)(F_{K,\sigma}^{n+1}(U_K^{n+1},U_{K,\sigma}^{n+1}) - F_{K,\sigma}^{n+1}(U_K^{n+1},U_K^{n+1}))\bar{U}_K^{n+1} \\ &\geq -\max(|A|,|B|)Lip(F_{K,\sigma}^{n+1})(B-A)Tm(\partial\Omega) \end{aligned}$$

And it is easy to see that

$$E2,2 - E2,2ext \geq -Lip(F_{K,\sigma}^{n+1})(B-A)\mathcal{N}_D(\bar{U}_D)(m(\Omega)T)^{\frac{1}{2}}$$

By collecting the previous inequalities, we complete the proof, the constant C depending on the constant of regularity , ξ, the measure of $m(\Omega)$, T, B, A, $Lip(F_{K,\sigma}^{n+1})$, $\|\bar{u}_t\|_{L^1(Q)}$ and $\|\nabla\bar{u}\|_{L^2(Q)}$
□

3.8 Continuous entropy inequalities

Let D be an admissible discretization of Q and let U_D be the discrete solution of the scheme (Proposition 3.6.1). We define a piecewise constant function $u_D$on Q, that we call an approximate solution, by

$$u_D(t,x) = U_K^{n+1} \qquad , \ t \in (t_n, t_{n+1}), \ x \in K. \tag{3.34}$$

From the discrete a priori estimates we made on U_D at Section 3.7 and Section 3.6 we deduce easily continuous estimates on u_D. By this way, we are able to prove the following theorem that is the key of the convergence proof.

Theorem 3.8.1 (Continuous approximate entropy inequalities) *Let D be an admissible discretization of Q. Let u_D be the corresponding approximate solution defined below. We have the following inequalities: $\forall \kappa \in \mathbb{R}$, $\forall \psi \in \mathcal{C}^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$, such that $\eta_\kappa^+(\bar{u})\psi = 0$ on Σ, we have*

$$\begin{aligned} \int_0^T \int_\Omega (\eta_\kappa^+(u_D)\psi_t + \Phi_\kappa^+(t,x,u_D)\cdot\nabla\psi + \eta_{\varphi(\kappa)}^+(\varphi(u_D))\Delta\psi) \quad &- \quad \int_\Sigma \eta_{\varphi(\kappa)}^+(\varphi(\bar{u}))\nabla\psi\cdot\mathbf{n} \\ &+ \quad \int_\Omega \eta_\kappa^+(u_0)\psi(0) \geq -\mathcal{E}_D^+(\varphi) \end{aligned} \tag{3.35}$$

and $\forall \psi \in \mathcal{C}^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$, such that $\eta_\kappa^-(\bar{u})\psi = 0$ on Σ, we have

$$\begin{aligned} \int_0^T \int_\Omega (\eta_\kappa^-(u_D)\psi_t + \Phi_\kappa^+(t,x,u_D)\cdot\nabla\psi + \eta_{\varphi(\kappa)}^-(\varphi(u_D))\Delta\psi) \quad &- \quad \int_\Sigma \eta_{\varphi(\kappa)}^-(\varphi(\bar{u}))\nabla\psi\cdot\mathbf{n} \\ &+ \quad \int_\Omega \eta_\kappa^-(u_0)\psi(0) \geq -\mathcal{E}_D^-(\varphi) \end{aligned} \tag{3.36}$$

where $\mathcal{E}_D^-$ and $\mathcal{E}_D^+$ tend to zero when the size of the discretization tends to zero.

Remark 3.8.1 *The condition $\eta_\kappa^+(\bar{u})\psi = 0$ is equivalent to $\psi = 0$ on the open subset $\{\bar{u} > \kappa\}$, and $\eta_\kappa^-(\bar{u})\psi = 0$ is equivalent to $\psi = 0$ on $\{\bar{u} < \kappa\}$. In particular, if $\bar{u} = 0$ and $\varphi(0) = 0$, we find the same formulation as Carrillo in [Car99].*

Remark 3.8.2 *Any entropy solution is a weak solution. Indeed, if we take tests functions nul on Σ, for $\kappa = min(u)$ we have that $\eta_\kappa^+(u) = u$.*

Proof. We will only prove Inequality (3.35), the proof of (3.36) being the same. Let $\kappa \in \mathbb{R}$ and $\psi \in \mathcal{C}^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$, a nonnegative function satisfying $\eta_\kappa^+(\bar{u})\psi = 0$ on Σ. We define discrete values of ψ as in the following:

$$
\begin{aligned}
&\Psi_K^0 = \psi(0, x_K), \quad K \in \mathcal{T}, \\
&\Psi_K^{n+1} = \frac{1}{\delta t_n} \int_{t^n}^{t^{n+1}} \psi(\cdot, x_K), \quad K \in \mathcal{T}, n \in [\![0, N]\!], \\
&\Psi_\sigma^{n+1} = \frac{1}{\delta t_n} \int_{t^n}^{t^{n+1}} \psi(\cdot, x_\sigma), \quad \sigma \in \mathcal{E}_{ext}, n \in [\![0, N]\!].
\end{aligned}
$$

In the same way as for U_D, we denote $\Psi_{K,\sigma}^{n+1} = \Psi_L^{n+1}$ if $\sigma = K|L$ and $\Psi_{K,\sigma}^{n+1} = \Psi_\sigma^{n+1}$ if $\sigma \in \mathcal{E}_{ext,K}$.

It is easy to verify on the Definition of $\Phi_{K,\sigma,\kappa}^{+,n+1}$ that it is a conservative flux consistent with Φ_κ^+ hence we have

$$
\sum_{\sigma \in \mathcal{E}_K} m(\sigma) \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_K^{n+1}) = 0, \qquad \forall K \in \mathcal{T}, n \in [\![0, N]\!].
$$

The discrete entropy inequality (3.29) can then be rewritten :

$$
\begin{aligned}
\frac{\eta_\kappa^+(U_K^{n+1}) - \eta_\kappa^+(U_K^n)}{\delta t^n} m(K) \quad &+ \sum_{\sigma \in \mathcal{E}_K} m(\sigma)(\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_K^{n+1})) \\
&- \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma(\eta_{\varphi(\kappa)}^+(\varphi(U_{K,\sigma}^{n+1})) - \eta_{\varphi(\kappa)}^+(\varphi(U_K^{n+1}))) \leq 0,
\end{aligned}
\tag{3.37}
$$

Let us multiply (3.37) by $\delta t_n \Psi_K^{n+1}$, and sum other $K \in \mathcal{T}$, and $n \in [\![0, N]\!]$. It holds

$$
T1 + T2 + T3 \leq 0,
$$

where

$$
T1 \quad = \quad \sum_{n=0}^{N} \sum_{K \in \mathcal{T}} m(K)(\eta_\kappa^+(U_K^{n+1}) - \eta_\kappa^+(U_K^n))\Psi_K^{n+1}
$$

by summing over the edges, $T2 = T2int + T2ext$, with

$$
\begin{aligned}
T2int = \sum_{n=0}^{N} \delta t_n \sum_{K \in \mathcal{T}} \frac{1}{2} \sum_{\sigma \in \mathcal{E}_{int,K}} m(\sigma) \quad &(\Psi_K^{n+1}(\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_K^{n+1})) \\
&- \Psi_{K,\sigma}^{n+1}(\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_{K,\sigma}^{n+1}, U_{K,\sigma}^{n+1})))
\end{aligned}
$$

and

$$
T2ext = \sum_{n=0}^{N} \delta t_n \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{ext,K}} m(\sigma) \Psi_K^{n+1}(\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_K^{n+1})).
$$

$T3 = T3int + T3ext$, with

$$T3int = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} \tau_\sigma(\eta^+_{\varphi(\kappa)}(\varphi(U_K^{n+1})) - \eta^+_{\varphi(\kappa)}(\varphi(U_{K,\sigma}^{n+1})))(\Psi_K^{n+1} - \Psi_{K,\sigma}^{n+1})$$

and

$$T3ext = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma(\eta^+_{\varphi(\kappa)}(\varphi(U_K^{n+1})) - \eta^+_{\varphi(\kappa)}(\varphi(U_{K,\sigma}^{n+1})))\Psi_K^{n+1}$$

We wish to prove that

$$I1 + I2 + I3 \leq \mathcal{E}_D^+(\psi)$$

with

$$\begin{array}{rcl} I1 & = & -\int_0^T\int_\Omega \eta^+_\kappa(u_D)\psi_t - \int_\Omega \eta^+_\kappa(u_0)\psi(0), \\ I2 & = & -\int_0^T\int_\Omega \Phi^+_\kappa(t,x,u_D)\cdot\nabla\psi, \\ I3 & = & -\int_0^T\int_\Omega \eta^+_{\varphi(\kappa)}(\varphi(u_D))\Delta\psi + \int_\Sigma \eta^+_{\varphi(\kappa)}(\varphi(\bar{u}))\nabla\psi\cdot\mathbf{n} \end{array}$$

To this purpose, we compare $I1$ to $T1$, $I2$ to $T2$ and $I3$ to $T3$.

1. Comparison of $I1$ and $T1$

$$I1 \quad = \quad \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} (\eta^+_\kappa(U_K^{n+1}) - \eta^+_\kappa(U_K^n)) \int_K \psi(t^{n+1},x) + \int_\Omega (\eta^+_\kappa(u_D^0) - \eta^+_\kappa(u_0))\psi(0,\cdot)$$

From the fact that η^+_κ is a Lipschitz continuous function with Lipschitz constant 1, we get

$$|I1 - T1| \leq \mathcal{E}^0_{\mathcal{T}}(\varphi) + \mathcal{E}^1_D(\varphi),$$

where

$$\mathcal{E}^0_{\mathcal{T}} = \int_\Omega |u_D^0 - u_0|\varphi(0,\cdot), \qquad \mathcal{E}^1_D(\varphi) = \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)|U_K^{n+1} - U_K^n| \int_{t^n}^{t^{n+1}} |\psi(t^{n+1},x) - \Psi_K^{n+1}|$$

2. Comparison of $I2$ and $T2$

In the same way as for $T2$, by using that u_D is piecewise constant and after a space integrate by parts, we obtain

$$I2 = I2int + I2ext$$

where $I2ext$ is the boundary term. By gathering by edges, we get

$$I2int = -\sum_{n=0}^{N}\sum_{K\in\mathcal{T}}\frac{1}{2}\sum_{\sigma\in\mathcal{E}_{int,K}}\Big(\int_{t^n}^{t^{n+1}}\int_\sigma \Phi_\kappa^+(\cdot,\cdot,U_K^{n+1})\cdot\mathbf{n}_{K,\sigma}\psi - \int_{t^n}^{t^{n+1}}\int_\sigma \Phi_\kappa^+(\cdot,\cdot,U_{K,\sigma}^{n+1})\cdot\mathbf{n}_{K,\sigma}\psi\Big),$$

and

$$I2ext = -\sum_{n=0}^{N}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_{ext,K}}\int_{t^n}^{t^{n+1}}\int_\sigma \Phi_\kappa^+(\cdot,\cdot,U_K^{n+1})\cdot\mathbf{n}_{K,\sigma}\psi$$

We first treat the interior terms. If we compare Φ_κ^+ on the interface σ to $\Phi_{K,\sigma,\kappa}^{+,n+1}$, we get

$$\begin{aligned}
I2int = -\sum_{n=0}^{N}\sum_{K\in\mathcal{T}}\frac{1}{2}\sum_{\sigma\in\mathcal{E}_{int,K}}\Big(&\int_{t^n}^{t^{n+1}}\int_\sigma (\Phi_\kappa^+(\cdot,\cdot,U_K^{n+1})\cdot\mathbf{n}_{K,\sigma} - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_K^{n+1}))\psi\\
&\int_{t^n}^{t^{n+1}}\int_\sigma (\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_K^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_{K,\sigma}^{n+1}))\psi\\
&-\int_{t^n}^{t^{n+1}}\int_\sigma (\Phi_\kappa^+(\cdot,\cdot,U_{K,\sigma}^{n+1})\cdot\mathbf{n}_{K,\sigma} - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_{K,\sigma}^{n+1},U_{K,\sigma}^{n+1}))\psi\\
&-\int_{t^n}^{t^{n+1}}\int_\sigma (\Phi_{K,\sigma,\kappa}^{+,n+1}(U_{K,\sigma}^{n+1},U_{K,\sigma}^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_{K,\sigma}^{n+1}))\psi\Big)
\end{aligned}$$

Then we get

$$|I2int - T2int| \le \mathcal{E}_D^{c,int}(\psi) + \mathcal{E}_D^{b,int}(\psi)$$

where

$$\begin{aligned}
\mathcal{E}_D^{c,int}(\psi) = \Big|\sum_{n=0}^{N}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_{int,K}}\int_{t^n}^{t^{n+1}}\int_\sigma &\Phi_\kappa^+(\cdot,\cdot,U_K^{n+1})\cdot\mathbf{n}_{K,\sigma} - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_K^{n+1})\psi\\
&+\Phi_\kappa^+(\cdot,\cdot,U_{K,\sigma}^{n+1})\cdot\mathbf{n}_{K,\sigma} - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_{K,\sigma}^{n+1},U_{K,\sigma}^{n+1})\psi\ \Big|
\end{aligned}$$

$$\begin{aligned}
\mathcal{E}_D^{b,int}(\psi) = \sum_{n=0}^{N}\delta t_n\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_{int,K}}&|\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_K^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_{K,\sigma}^{n+1})|\Big|\int_{t^n}^{t^{n+1}}\int_\sigma \Psi_K^{n+1} - \psi\ \Big|\\
&+|\Phi_{K,\sigma,\kappa}^{+,n+1}(U_{K,\sigma}^{n+1},U_{K,\sigma}^{n+1}) - \Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1},U_{K,\sigma}^{n+1})|\Big|\int_{t^n}^{t^{n+1}}\int_\sigma \Psi_{K,\sigma}^{n+1} - \psi\ \Big|
\end{aligned}$$

Let us now study $I2ext$ and $T2ext$.

$$T2ext - I2ext \leq \mathcal{E}_D^{c,ext} + \mathcal{E}_D^{b,ext} + T_D^{b,ext}$$

where

$$\mathcal{E}_D^{c,ext}(\psi) = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \Big| \int_{t^n}^{t^{n+1}} \int_{\sigma} (\Psi_K^{n+1} - \psi)\Phi_\kappa^{+}(\cdot,\cdot,U_K^{n+1}) \cdot \mathbf{n}_{K,\sigma} \Big|$$

$$\mathcal{E}_D^{b,ext}(\psi) = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} m(\sigma) |(\Psi_K^{n+1} - \Psi_\sigma^{n+1})\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1})|$$

$$T_D^{b,ext}(\psi) = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} m(\sigma) \Psi_\sigma^{n+1}\Phi_{K,\sigma,\kappa}^{+,n+1}(U_K^{n+1}, U_{K,\sigma}^{n+1})$$

But by Definition of $\Phi_{K,\sigma,\kappa}^{+,n+1}$, and from the monotony of the scheme, we have

$$\Phi_{K,\sigma,\kappa}^{+,n+1}(a,b) = F_{K,\sigma}^{n+1}(a\top\kappa, b\top\kappa) - F_{K,\sigma}^{n+1}(\kappa,\kappa) \geq -Lip(F_{K,\sigma}^{n+1})(b-\kappa)^{+}$$

So

$$T_D^{b,ext}(\psi) \quad \leq \quad \mathcal{E}_D^{b,ext}(\psi) = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} m(\sigma) Lip(F_{K,\sigma}^{n+1})(\bar{U}_\sigma^{n+1} - \kappa)^{+}\Psi_\sigma^{n+1},$$

Let us recall now that $\eta_\kappa^{+}(\bar{u})\psi = 0$ on Σ. Since $\bar{u}$ is uniformly continuous on Σ and ψ is a Lipschitz continuous function, and that the two functions are nonnegative, we have

$$\eta_\kappa^{+}(\bar{U}_\sigma^{n+1})\Psi_\sigma^{n+1} \leq Lip(\psi)(\delta t_n + diam(\sigma))M_{\bar{u}}(\delta t_n + diam(\sigma))$$

where $M_{\bar{u}}$ denotes the modulus of continuity of $\bar{u}$. And

$$\mathcal{E}_D^{b,ext}(\psi) \leq \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} m(\sigma) Lip(\psi)(\delta t_n + diam(\sigma))M_{\bar{u}}(\delta t_n + diam(\sigma))$$

2.Comparison of $I3$ and $T3$

In the same way as for $I2$ and $T2$, we compare separately the interior terms and the exterior terms. By using the Definition of u_D, we have

$$I3int = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} (\eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K,\sigma})) - \eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K})) \int_{t^n}^{t^{n+1}} \int_{\sigma} \nabla\psi\cdot\mathbf{n}_{K,\sigma}$$

So

$$|I3int - T3int| \leq \mathcal{E}^{c,int}_{3,D}(\psi)$$

where

$$\mathcal{E}^{c,int}_{3,D}(\psi) = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{\sigma\in\mathcal{E}_{int,K}} |\eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K,\sigma})) - \eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K})| \ | \int_{\sigma} \nabla\psi\cdot\mathbf{n}_{K,\sigma} - \tau_\sigma(\Psi^{n+1}_{K,\sigma} - \Psi^{n+1}_{K})|$$

Let us now treat $I3ext$ and $T3ext$.

$$\begin{aligned} I3ext - T3ext \quad &= \quad -\sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} (\ \eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K})) \int_{t^n}^{t^{n+1}} \int_{\sigma} \nabla\psi\cdot\mathbf{n} + 2\int_{t^n}^{t^{n+1}} \int_{\sigma} \eta^{+}_{\varphi(\kappa)}(\bar{u})\nabla\psi\cdot\mathbf{n} \\ &\quad -\tau_\sigma(\eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K})) - \eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K,\sigma}))\Psi^{n+1}_{K}\ ) \\ &\leq \quad T^{b,ext}_{3,D}(\psi) + \mathcal{E}^{c,ext}_{3,D} + \mathcal{E}^{bb,ext}_{3,D} + \mathcal{E}^{bbb,ext}_{3,D} \end{aligned}$$

where (using the fact η^{+}_{κ} is a Lipschitz continuous function with Lipschitz constant 1) we have

$$T^{b,ext}_{3,D}(\psi) = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma(\eta^{+}_{\varphi(\kappa)}(\varphi(\bar{U}^{n+1}_{\sigma})) - \eta^{+}_{\varphi(\kappa)}(\varphi(U^{n+1}_{K}))\Psi^{n+1}_{K})$$

$$\mathcal{E}^{c,ext}_{3,D}(\psi) = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} 2 \max_{u\in[A,B]} \eta^{+}_{\varphi(\kappa)}(\varphi(u))| \int_{t^n}^{t^{n+1}} \int_{\sigma} (\nabla\psi\cdot\mathbf{n} - \frac{\Psi^{n+1}_{\sigma} - \Psi^{n+1}_{K}}{d_K,\sigma})|$$

$$\mathcal{E}^{bb,ext}_{3,D}(\psi) = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \int_{t^n}^{t^{n+1}} \int_{\sigma} |\varphi(\bar{u}) - \varphi(\bar{U}^{n+1}_{\sigma})||\nabla\psi\cdot\mathbf{n}|$$

$$\mathcal{E}^{bbb,ext}_{3,D}(\psi) = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \int_{t^n}^{t^{n+1}} \int_{\sigma} |\varphi(\bar{U}^{n+1}_{\sigma}) - \varphi(U^{n+1}_{K})||\nabla\psi\cdot\mathbf{n}|.$$

Now using the fact that ψ is a nonnegative function, and the convexity of $\eta^+_{\varphi(\kappa)}$ we get

$$T^{b,ext}_{3,D}(\psi) \leq \mathcal{E}^{b,ext}_{3,D}(\psi) = \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma {\eta^+_{\varphi(\kappa)}}'(\varphi(\bar{U}^{n+1}_\sigma))\Psi^{n+1}_K(\varphi(\bar{U}^{n+1}_\sigma) - \varphi(U^{n+1}_K))$$

By the same argument as we used above, we have

$${\eta^+_{\varphi(\kappa)}}'(\varphi(\bar{U}^{n+1}_\sigma)\eta^+_{\varphi(\kappa)}(\varphi(\bar{U}^{n+1}_\sigma) \leq Lip(\psi)(\delta t_n + diam(\sigma))$$

so we obtain

$$|\mathcal{E}^{b,ext}_{3,D}(\psi)| \leq \sum_{n=0}^{N} \delta t_n \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_{ext,K}} \tau_\sigma Lip(\psi)(\delta t_n + diam(\sigma))|\varphi(\bar{U}^{n+1}_\sigma) - \varphi(U^{n+1}_K)|$$

The term on the right hand side of this inequality does not tend to zero in all cases. A simple way to ensure that it converges to zero is to suppose an uniform C.F.L. condition (with a C.F.L. number that can be very large).
It remains to estimate $\mathcal{E}^0_{\mathcal{T}}(\psi)$, $\mathcal{E}^1_D(\psi)$, $\mathcal{E}^{c,ext}_D(\psi)$, $\mathcal{E}^{cc,ext}_{2,D}(\psi)$, $\mathcal{E}^{c,ext}_{3,D}(\psi)$, $\mathcal{E}^{b,ext}_{3,D}(\psi)$, $\mathcal{E}^{bb,ext}_{3,D}(\psi)$ and $\mathcal{E}^{bbb,ext}_{3,D}(\psi)$. We verify that each of these terms tend to zero using energy estimates, weak BV inequality and regularity of the mesh.
□

3.9 Convergence of the scheme

Let D_n be a sequence of discretizations, such that $size(D_n)$ tends to zero. We wish to prove the convergence of u_{D_n} to an entropy solution of Problem (3.1). For that, using the uniqueness theorem 3.4.1, it suffices to show that up to a subsequence, u_{D_n} tends in the non-linear weak star sense to a entropy process solution of (3.1). We obtain compactness properties using estimates on u_{D_n} derived from discrete estimates on U_{D_n}. Then, it is not difficult to pass to the limit in Inequalities (3.35) and (3.36).

3.9.1 Non-linear weak star compactness

By using the maximum principle, u_{D_n} is bounded in $L^\infty(Q)$, so there exist $u \in L^\infty(Q\times(0,1))$, such that up to a subsequence, u_{D_n} tends to u in the non-linear weak star sense.

3.9.2 Compactness in $L^2(Q)$

From discrete estimates obtained in Proposition 3.7.1 we easily deduce the following inequalities on $z_D = \zeta(u_D) - \zeta(\bar{u}_D)$:

Proposition 3.9.1 (Space translate estimates) *There exist C_1 such that*

$$\forall y \in \mathbb{R}^d, \int_0^T \int_{\Omega_\xi} (z_D(t, x+y) - z_D(t,x))^2 dx dt \le C_1 |y|(|y| + \mathrm{size}(\mathcal{T})),$$

where $\Omega_y = \{x \in \Omega, [x, x+y] \subset \Omega\}$

By Hypothesis (H5), we have $\bar{u}_t \in L^1(Q)$. By using that U_D satisfy a discrete evolution Equation (3.26) and if we recall that u_D is linked to U_D by Formula (3.34), we can also deduce the following time translate estimate on z_D:

Proposition 3.9.2 (Time translate estimates) *There exist C_2 such that*

$$\forall s > 0, \int_0^{T-s} \int_\Omega (z_D(t+s, x) - z_D(t,x))^2 dx dt \le C_2 \ s$$

Let us recall the following compactness theorem.

Theorem 3.9.1 (Fréchet-Kolmogorov's theorem) *Let Q be an open bounded subset of $\mathbb{R}^k$ and $(u_n)_{n \in \mathbb{N}}$ be a bounded sequence in $L^2(\mathbb{R}^k)$ such that*

$$\lim_{|\delta| \to 0} \sup_{n \in \mathbb{N}} \|u_n(\cdot + \delta) - u_n(\cdot)\|_{L^2(Q)} = 0,$$

then there exists $u \in L^2(Q)$ such that, up to a subsequence,

$$u_n \to u \text{ for the strong topology of } L^2(Q) \text{ as } n \longrightarrow \infty.$$

Since z_D vanishes on Σ, we can extend z_d by zero out of Q, without any problem. Then, we can use Kolmogorov theorem that gives (with an argument used many times for parabolic problems) that there exist $z \in L^2(0, T, H^1(\Omega))$, such that up to a subsequence, $z_{D_n} \to z$ in $L^2(Q)$.

Now, we recall that $z_D = \zeta(u_d) - \zeta(\bar{u}_d)$, and we remark that $\zeta(\bar{u}_d)$ converges to $\zeta(\bar{u})$ in $L^2(Q)$ so $\zeta(u_{D_n})$ tends to $\zeta(\bar{u}) + z$. From the non-linear weak star convergence, $\zeta(u_{D_n})$ converges also to $\zeta(u)$ weakly in $L^\infty(Q)$, so $\zeta(\bar{u}) + z = \zeta(u)$. In particular, $\zeta(u)$ does not depend on the last argument α and $\zeta(u) = \zeta(\bar{u})$ a.e. on Σ.

As we said in the preceding, it remains to take the limit in the continuous entropy inequalities to obtain that u is an entropy process solution. Then by the uniqueness theorem 3.4.1, u does not depend on α and is the unique solution of Problem (3.1). Besides all the sequence u_{D_n} is convergent (u is the unique possible limit). Moreover by definition of the nonlinear weak star convergence $(u_{D_n})^2$ also converges to $(u)^2$. So u_{D_n} converges to u in $L^2(Q)$, and in all $L^p(Q)$, for $1 \le p < +\infty$. The proof is complete.

Chapitre 4

Résultats numériques

4.1 Test parabolique dégénéré 1D

Les données utilisées pour ce premier test sont celles données par S. Evje et K.H. Karlsen dans [EK00b]. Cet exemple est instructif aussi bien du point vue de la convergence du schéma que pour la compréhension des phénomènes de dégénérescence.

$$\Omega = [0,1], \bar{u}(0) = 1, \bar{u}(1) = 0, f(u) = \frac{1}{8}u^2,$$

$$\varphi(u) = \begin{cases} 0 & \text{si} \quad 0 \leq u < 0.5 \\ 0.125(u-0.5)^2 & \text{si} \quad 0.5 \leq u < 0.6 \\ 0.125(0.01 + 0.2(u-0.6)) & \text{si} \quad 0.6 \leq u \leq 1 \end{cases}$$

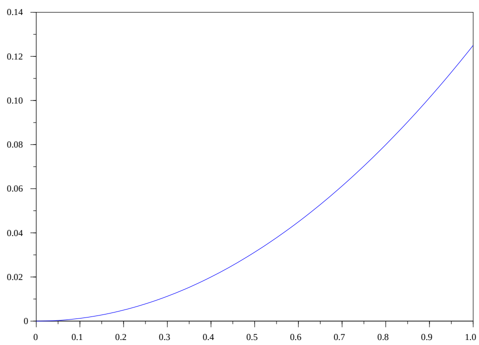


Fig 3. *Graphe de la fonction f.*

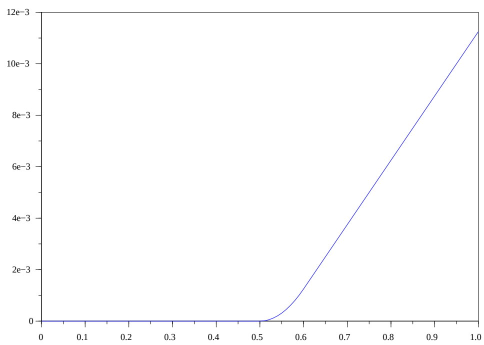


Fig 4. *Graphe de la fonction φ.*

Pour observer le comportement du schéma et avoir une bonne approximation de la solution exacte, nous avons fait tourner l'algorithme avec différents pas d'espaces, de manière à diminuer progressivement la taille des cellules .

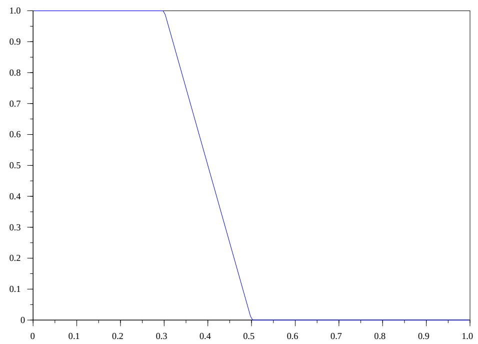


Fig 5. *Condition initiale* u_0.

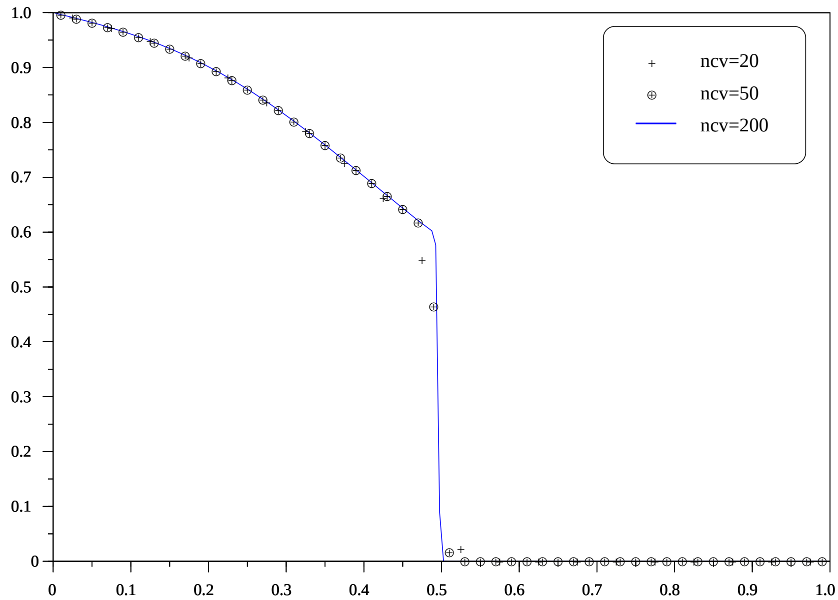


Fig 6. *Résultats au temps t=0.1 pour différents maillages.*

Plusieurs enseignements sont à retirer de ce test. Tout d'abord, on voit très bien le profil de la solution exacte apparaître lorsqu'on augmente le nombre de points de discrétisation. On peut séparer le domaine en deux zones. Dans la première zone, coïncidant avec l'ensemble $\{u > 0.5\}$, la solution, dont la dérivée présentait une discontinuité au point d'abscisse $x = 0.4$ à l'instant initial , est régularisée par le terme de diffusion. Dans l'autre zone $\{u < 0.5\}$ la solution présente un profil de choc classique pour une équation de type Burgers. Ensuite, la résolution en 1D est bonne avec très peu de points, même si le schéma utilisé est un schéma d'ordre 1 et le choc est bien résolu avec seulement 20 points. Avec 50 points, l'approximation est presque aussi bonne que pour 200 points. En particulier, le positionnement de la discontinuité ne nécessite que peu de points de discrétisation.

Techniquement le schéma a été programmé avec le logiciel libre Scilab 5.2 distribué par l'INRIA. Les résultats proviennent du schéma volumes finis décentré amont explicite en temps, qui pose bien entendu un peu moins de difficultés de programmation et s'avère aussi efficace sur des cas aussi simples.

4.2 Test parabolique dégénéré 2D

Nous nous sommes intéressés à un cas similaire au cas présenté au paragraphe 4.1 en 2D cette fois ci. Le test en deux dimensions d'espace permet de mieux comprendre comment les phénomènes de choc et de détente sont mêlés aux phénomènes de diffusion.
Toutes les conditions du chapitre 2 sont respectées. La fonction φ est définie par $\varphi(u) = 0.005 * (u - 0.5)^+$.
Le champ de vitesse q tangent au bord du domaine est défini par

$$q(x, y) = (10(2x - 1)(y^2 - y), -10(x^2 - x)(2y - 1)).$$

La fonction φ est donc lipschitzienne et croissante au sens large sur l'intervalle $[0, 1]$. Elle est constante sur l'intervalle $[0, 0.5]$ et strictement croissante sur l'intervalle $[0.5, 1]$. La seule différence avec la fonction φ du paragraphe précédent est sa régularité, mais en théorie comme en pratique, seul le caractère lipschitzien est utilisé. L'algorithme de Newton fonctionne correctement malgré les sauts de dérivée. Deux exemples de condition initiale et de conditions limites sur le bord ont retenu notre attention.

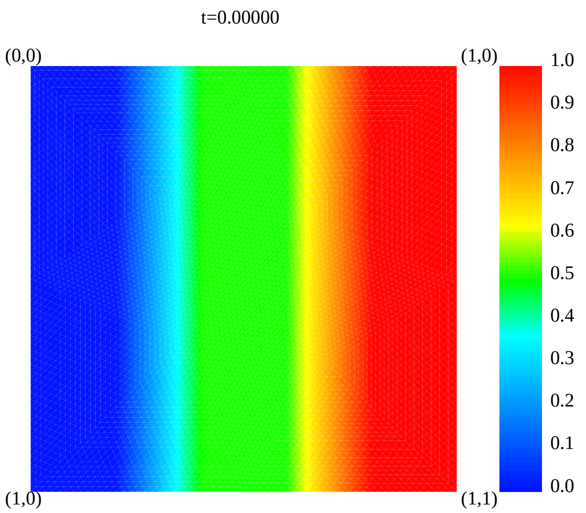
Le premier exemple a déjà été présenté au chapitre 2. Nous avons mis en lumière, dans ce premier exemple, l'effet de la dégénérescence du terme de diffusion sur une plage, en observant l'évolution de deux carrés identiques mais avec des valeurs de saturations différentes.

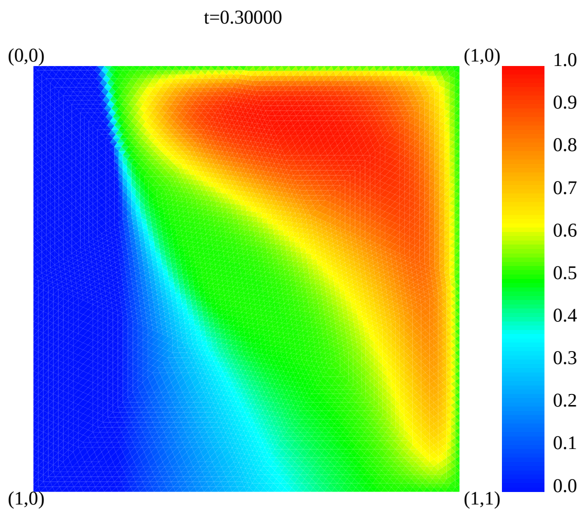
Pour le deuxième exemple présenté ici, le phénomène est un peu plus complexe. Les fonctions utilisées sont les suivantes :

$$f(u) = \frac{u^2}{2},\ \bar{u} = 0,\ u_0(x, y) = g(x)$$

où g est définie comme suit :

$$g(x) = \begin{cases} 0 & \text{si} \quad 0 \leq x < 0.2 \\ x - 0.2 & \text{si} \quad 0.2 \leq x < 0.4 \\ 0.5 & \text{si} \quad 0.4 \leq x < 0.6 \\ (x - 0.6) + 0.5 & \text{si} \quad 0.6 \leq x < 0.8 \\ 1 & \text{si} \quad 0.8 \leq x \leq 1 \end{cases}$$





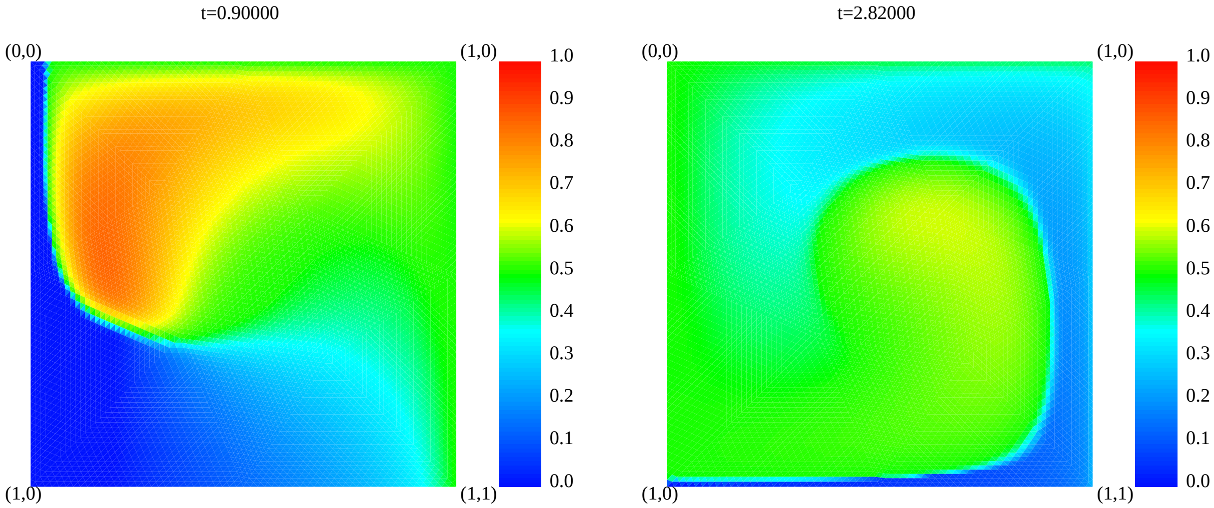


Fig 7. *Solution du problème parabolique dégénéré à différents instants*

On peut faire les observations suivantes :

1. Si la diffusion est nulle, le problème correspond à une simple équation de Burgers, qui est le cas le plus simple d'équation hyperbolique non linéaire. Entre deux zones où la valeur de la saturation est différente et inférieure à 0.5, il se forme un choc ou une détente selon que la valeur de la saturation est plus grande en amont ou en aval, l'amont et l'aval étant définis par le champ de vitesse $\mathbf{q}$. On observe, sur le demi-plan supérieur de la figure, un net resserrement des lignes de niveau, ce qui correspond à la formation d'un choc. Au contraire on observe un étalement dans la demi-partie inférieure et il s'agit là d'une onde de détente. Ce phénomène est restreint précisément à la zone où la diffusion s'annule, c'est à dire la zone $\{u < 0.5\}$ (couleur bleue et verte).

2. Pour des valeurs de la saturation plus grandes que 0.5, le phénomène change de nature. On observe qu'il n'y a plus ni choc ni détente dans la partie joignant les zones vertes et rouges. La zone $\{u \geq 0.5\}$ est transportée par le champ de vitesse $\mathbf{q}$ en moyenne, comme le reste du fluide, mais son évolution est identique à ce qu'on pourrait observer pour l'équation de la chaleur : il y a un étalement progressif vers la saturation moyenne, ce qui correspond sur la figure à une progressive évolution vers une couleur jaune puis verte de la zone initialement à dominante rouge.

3. Les conditions au bord imposées à la solution approchée par le schéma correspondent exactement à la condition $\varphi(u) = \varphi(\bar{u})$ imposée au sens fort dans la définition d'une solution faible (cf (1.8),(2.7) ou Definition 3.3.1). Imposer cette condition ici revient en fait à imposer la condition $\bar{u} \in [0; 0.5]$. Ainsi pour les valeurs de u inférieures à 0.5, cette condition est toujours vérifiée quelle que soit la trace de la solution. On peut considérer dans ce cas que le domaine est artificiel, au sens où le bord ne modifie ni le flux ni la valeur au bord imposée. On observe très bien ceci sur les figures ci-dessus, où aucun changement de la valeur de la saturation n'intervient près du bord si $u \in [0; 0.5]$, tandis que lorsque $u > 0.5$, c'est la condition $u = 0.5$ qui est imposée de part la régularité de la solution (frange verte entre le bord et la zone à dominante rouge).

Techniquement, les figures ont été obtenues par le schéma volumes finis décentré amont, implicite en temps. Ce schéma nécessite de calculer approximativement une racine pour un système d'équations non linéaires avec autant d'équations qu'il y a de volumes de contrôle (12600 sur l'exemple présenté). Nous avons fait le choix d'utiliser une méthode de Newton avec une résolution directe du système linéaire par une procédure de Gauss par bande. La résolution ne pose pas de difficulté, une approximation de la racine étant obtenue en moins de 5 itérations (pour des pas de temps raisonnables). Le maillage utilisé est construit à partir d'une grille grossière de triangles représentée ci dessous pour 2 domaines différents. Pour respecter les hypothèses de régularité sur le maillage dont nous avons besoin pour démontrer les principaux théorèmes de convergence (Chapitres 1, 2 et 3), il suffit d'imposer aux triangles des angles strictement inférieurs à $\frac{\pi}{2}$. Ensuite, chacun des triangles est raffiné pour obtenir un pavage par des triangles homothétiques plus petits. Ainsi, on diminue la taille des volumes de contrôle tout en gardant la propriété de régularité et d'admissibilité du maillage, le centre du cercle circonscrit étant pris comme centre de maille.

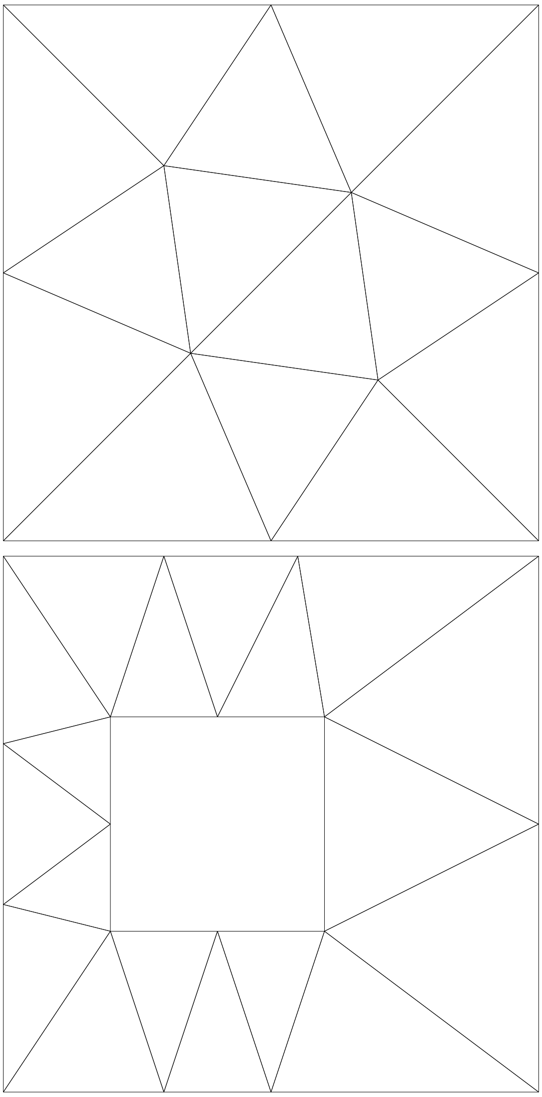


Fig 8. *Maillage grossier*

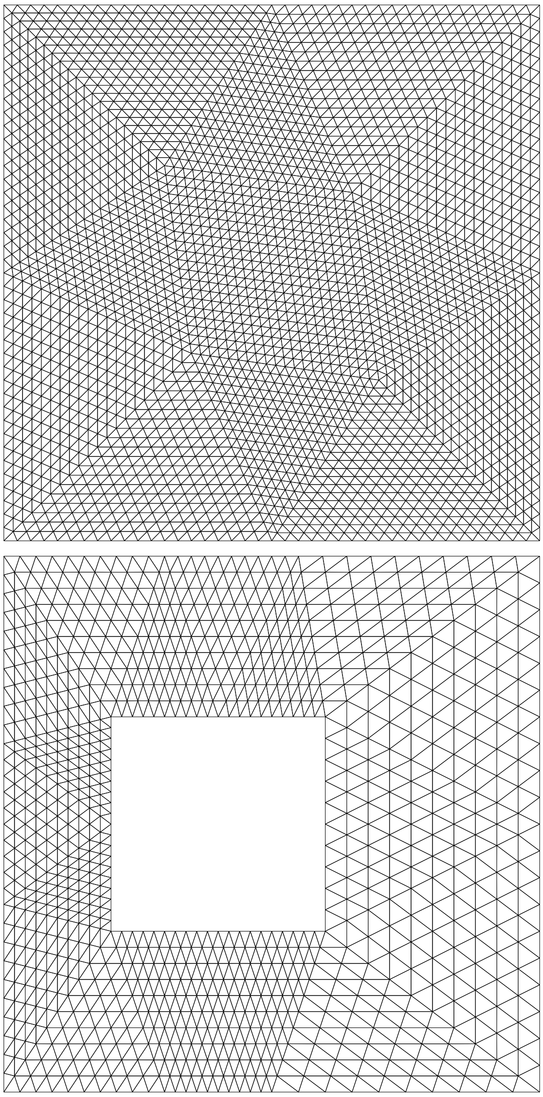


Fig 9. *Maillage raffiné*

Ce type de maillage permet facilement de produire un schéma pour des domaines qui peuvent être non convexes (quelques exemples sont donnés en illustration ci dessus). Nous avons programmé l'ensemble des procédures en langage C, sous Unix. Les sorties graphiques utilisent le logiciel de dessin vectoriel Xfig 3.2 faisant partie des programmes de base fournis avec le système UNIX.

Partie II

Analyse numérique d'un modèle simplifié d'écoulement diphasique incompressible en milieu poreux

Introduction

Cette partie est consacrée à l'étude d'un modèle simplifié d'écoulement diphasique en milieu poreux avec deux phases incompressibles et non miscibles. Pour l'étude mathématique, nous avons supposé que le milieu était homogène isotrope et les effets de la gravité sont négligés. Physiquement, ce modèle est acceptable pour l'étude de milieux constitués d'une seule roche sans direction privilégiée, comme l'argile ou le sable, ou pour un milieu horizontal. Mathématiquement, ces hypothèses enlèvent les difficultés liées à la non homogénéité du milieu. Nous pouvons ainsi nous concentrer sur l'influence de la pression capillaire. Ce choix justifie également l'étude approfondie des équations paraboliques dégénérées faite dans la première partie, la fonction φ étant en étroit lien avec la pression capillaire par la formule

$$\varphi'(u) = -\frac{k_1(u)k_2(u)}{k_1(u) + k_2(u)}{p_c}'(u) \tag{4.1}$$

Dans le préambule qui suit est expliquée l'origine des équations étudiées. En particulier, il est expliqué précisément où interviennent les simplifications dans le modèle envisagé. Ce paragraphe est destiné à mettre en évidence la signification des grandeurs mises en jeu, leur ordre de grandeur et les unités dans lesquelles elles s'expriment. Le principal objectif étant de montrer la cohérence des hypothèses que nous faisons sur les différentes fonctions intervenant dans le système couplé étudié.
A partir de ce modèle simplifié, on peut manipuler les équations à la manière de Chavent et Jaffré [CJ86] pour se ramener à un problème couplé parabolique elliptique. Grâce aux travaux antérieurs sur la discrétisation des équation aux dérivées partielles en volumes finis [EGH00b], il est assez simple de construire un schéma pour ce type d'équations qui respecte les principes du maximum sur la saturation et qui permette d'obtenir d'autres estimations a priori sur la pression et sur la saturation. L'étude mathématique de la convergence de ce schéma, que l'on a nommé le schéma "des mathématiciens", a fait l'objet d'un article soumis début 2001 [Mic01].
Le deuxième chapitre est constitué lui aussi d'un article qui devrait être soumis très prochainement. Il est consacré à un schéma appelé schéma des pétroliers qui consiste à discrétiser les deux équations de conservation directement en effectuant un décentrage amont sur chacune des phases séparément. Ce schéma a de nombreux avantages. Le premier est de n'utiliser que des fonctions physiques. C'est à dire que toutes les fonctions utilisées sont mesurables ou sont des fonctions simples de fonctions mesurables expérimentalement. Comme pour le schéma du chapitre 1, on prouve que le principe du maximum est respecté. Puis, sous des hypothèses un peu plus restrictives mais toutefois encore acceptables au niveau physique, on montre la convergence du schéma en passant par des estimations originales. Les preuves sont beaucoup plus difficiles que pour le schéma des mathématiciens, à cause de la présence de deux décentrages différents. Cet article est également motivé par le fait que le schéma étudié est utilisé en pratique dans l'ingénierie pétrolière.
Ces schémas ont fait l'objet de tests numériques avec les données physiques décrites dans le préambule. Les résultats figurent dans le dernier chapitre de cette deuxième partie.

Justification du modèle étudié

Pour comprendre le modèle d'écoulement diphasique incompressible que nous considérons, il faut se placer au niveau macroscopique, c'est à dire à une échelle où les propriétés d'une même roche sont supposées continues. Elles sont alors représentées par des fonctions de densité.
La première de ces propriétés est la porosité. On la note habituellement Φ. Si la roche n'est pas déformable dans les conditions expérimentales considérées, cette porosité ne dépendra que de l'endroit où on se place, autrement dit $\Phi = \Phi(x)$. La porosité est la proportion occupée par le volume des pores qui permettent l'écoulement. C'est donc une quantité strictement comprise entre 0 et 1 et elle se mesure à l'aide d'un porosimètre à mercure.
Le milieu est ensuite caractérisé par une matrice de perméabilité $K(x)$. Si le milieu est isotrope, c'est à dire si aucune direction particulière n'est privilégiée, $K(x) = k(x)I$, où $k(x) \in \mathbb{R}$. La perméabilité est une mesure des efforts necessaires pour faire traverser le milieu poreux au fluide. Par exemple k sera plus petit pour l'argile que pour du sable (cf Tableau 1). On suppose d'autre part que le milieu est homogène, c'est à dire qu'en tout point on a les mêmes caractéristiques, ainsi $k(x) = k$ et $\Phi(x) = \Phi$.
Les équations qui régissent les écoulements en milieu poreux proviennent de la loi de Darcy qui s'énonce comme suit : soit $\vec{v}$ la vitesse apparente d'un fluide dans un milieu poreux de perméabilté K, alors

$$\vec{v} = -\frac{K}{\mu}[\nabla p - \rho g z]$$

où μ est la viscosité du fluide, p sa pression, ρ sa densité volumique et g la gravité. Nous négligeons les forces de gravité, ce qui sous entend que dans le mileiu poreux étudié les déplacements se font horizontalement. Ces termes devront être ajoutés pour les simulations numériques sur les cas réels.
Lorsque deux phases sont présentes en même temps, une loi similaire est valable, mais il faut ajouter un coefficient de perméabilité relative, qui est nécessaire pour représenter les interactions dues à la présence simultanée des deux fluides. Le modèle se présente alors sous la forme de deux équations de conservation de la masse couplées :

$$(\rho_1 \Phi \alpha_1)_t + \text{div}(-\frac{\rho_2 \tilde{k}_1}{\mu_1} \nabla p_1) = Q_1$$
$$(\rho_2 \Phi \alpha_2) + \text{div}(-\frac{\rho_2 \tilde{k}_2}{\mu_2} \nabla p_2) = Q_2$$

où α_1 est la saturation de la première phase (proportion de volume poreux occupé par la première phase) , α_2 la proportion de la seconde phase, p_1 et p_2 les pressions respectives des deux phases. Q_1 le débit volumique en espèce de la première phase et Q_2 le débit volumique en espèces de la seconde phase. La différence entre les deux pressions est appelée pression capillaire. On note $p_c = p_2 - p_1$. Cette pression capillaire est l'expression au niveau macroscopique des forces de tension superficielle présentes sur l'interface entre deux fluides. On prend généralement pour inconnues principales $u = \alpha_1$, la saturation de la première phase et $p = p_1$ la pression cette même phase. On suppose également que la pression capillaire ne dépend que de la saturation u, autrement dit $p_c = p_c(u)$. De part sa définition, pc est clairement une fonction décroissante dont l'allure est proche de celle représentée à la figure 10 . Si les sources ont une distribution volumique, après avoir divisé les équations par les constantes de porosité, de perméabilité et de densité volumique, le système se réécrit finalement sous la forme simple suivante :

$$u_t - \operatorname{div}(k_1(u)\nabla p) = f(c)\,\overline{s} - f(u)\,\underline{s} \tag{4.2}$$
$$u_t - \operatorname{div}(k_2(u)\nabla(p + p_c(u))) = (1 - f(c))\,\overline{s} - (1 - f(u))\,\underline{s} \tag{4.3}$$

où $\overline{s}$ représente un terme la source au puits injecteur et $\underline{s}$ un terme source au puits producteur, c étant la concentration en injection. Les fonctions k_1 et k_2 sont définies par

$$k_1(u) = \frac{\tilde{k}_1(u)}{\mu_1 \Phi}, \qquad k_2 = \frac{\tilde{k}_2(u)}{\mu_2 \Phi}$$

et sont appelées mobilités réduites. La fonction f (the "fractionnal flow") est définie par

$$f(u) = \frac{k_1(u)}{k_1(u) + k_2(u)}.$$

La propriété de monotonie de p_c et son comportement est essentiel dans les preuves des théorèmes de convergence énoncés aux chapitres 5 et 6. Les fonctions k_1 et k_2 sont quant à elles respectivement croissantes et décroissantes et la somme $M = k_1 + k_2$ est toujours minorée par un réel strictement positif. La fonction f est donc croissante et comprise entre 0 et 1, comme c'est le cas pour l'exemple de la figure 11.

Le tableau ci dessous regroupe les données physiques approximatives que l'on pourrait rencontrer dans un exemple concret.

Porosité sable	Φ_s	0.2	
Porosité argile	Φ_a	0.1	
Perméabilité sable	k_s	10^{-12}	m^2
Perméabilité argile	k_a	10^{-15}	m^2
Densité Vol. Eau	ρ_w	1000	$kg.m^{-3}$
Densité Vol. Air	ρ_a	3	$kg.m^{-3}$
Densité Vol. Huile	ρ_o	900	$kg.m^{-3}$
Viscosité Eau	μ_w	10^{-3}	$Pa.s$
Viscosité Air	μ_a	10^{-5}	$Pa.s$
Viscosité Huile	μ_o	$10^{-3} \to 10^{-1}$	$Pa.s$

Tableau 1. *Données physiques*

Comme au chapitre 5, le système (4.2)(4.3) se réécrit sous la forme d'un système couplé formé d'une équation parabolique hyperbolique sur la saturation et d'une équation elliptique sur la pression.Le terme de diffusion qui intervient dans l'équation parabolique en saturation est $-\Delta\varphi(u)$ où φ est définie par (4.1) et le terme de convection est $\operatorname{div}(qf(u))$ où $\mathbf{q}$ est la somme des vélocités des deux phases :

$$\mathbf{q} = -k_1(u)\nabla p - k_2(u)\nabla(p + p_c(u)). \tag{4.4}$$

Chapitre 5

Convergence du schéma volumes finis des "mathématiciens"

Résumé

Dans ce chapitre, nous montrons la convergence d'un schéma volumes finis vers une solution faible du problème couplé d'équations aux dérivées partielles sur la saturation $u = u(x,t)$ et la pression globale $\theta = \theta(x,t)$ suivant :

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta\varphi(u) - \operatorname{div}(f(u)M(u)\nabla\theta) &= c\,\overline{s} - u\,\underline{s}, \\ -\operatorname{div}(M(u)\nabla\theta) &= \overline{s} - \underline{s}, \end{aligned}$$

Le domaine étudié est le domaine cylindrique $\Omega \times (0,T)$. La preuve s'appuie principalement sur des estimations a priori pour u dans $L^2(\Omega \times (0,T))$, $\varphi(u)$ dans $L^2(0,T,H^1(\Omega))$ et θ dans $L^2(0,T,H^1(\Omega))$. Nous nous sommes placés sous l'hypothèse de faible dégénérescence de φ, c'est à dire dans le cas où φ est continue et strictement croissante. Dans ce cas, nous montrons une convergence forte pour la saturation et une convergence faible pour la pression.

Cette méthode repose sur la structure elliptique parabolique presque découplée du système. Les preuves données dans cet article fonctionnent en effet presque uniquement grâce aux idées utilisées pour les preuves de convergence en elliptique ou en parabolique hyperbolique. La fin de la preuve de convergence est néanmoins originale et elle fait appel à une technique de régularisation qui a été utile également dans le "liminf lemma" 2.5.2 au chapitre 2. Malheureusement, le schéma a montré des faiblesses au cours de la mise en oeuvre numérique. Ceci n'est pas très visible au niveau des résultats, mais le coût en programmation est largement supérieur au schéma présenté au chapitre 6, notamment parce que la fonction φ n'est pas une fonction simple de fonctions physiques données. Les résultats numériques présentés dans cet article ont été effectués en maillage rectangulaire.

A finite volume scheme for the simulation of two-phase incompressible flow in porous media.

Abstract

We prove the convergence of a finite volume approximation to a weak solution of the following coupled system with two unknowns $u = u(x,t)$ and $\theta = \theta(x,t)$:

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta\varphi(u) - \operatorname{div}(f(u)M(u)\nabla\theta) &= c\overline{s} - u\underline{s}, \\ -\operatorname{div}(M(u)\nabla\theta) &= \overline{s} - \underline{s}, \end{aligned}$$

on a bounded cylindrical domain $\Omega \times (0,T)$ with Neumann boundary conditions where the notations are given in the course of the chapter. The proof is mainly based on discrete *a priori* estimates for u in $L^\infty((0,T)\times\Omega)$, for $\varphi(u)$ in $L^2(0,T,H^1(\Omega))$ and for $\theta \in L^\infty(0,T,H^1(\Omega))$. Strong convergence of the approximate solution for u is obtained under the assumption that φ is a strictly nondecreasing continuous function.

5.1 Introduction

A simple model of displacement of two incompressible fluids in a porous media (see [Bea67]) with volumetric source and sink term is given by

$$\frac{\partial u}{\partial t} - \operatorname{div}(k_1(u)\nabla p) = c\overline{s} - u\underline{s}, \tag{5.1}$$

$$\frac{\partial(1-u)}{\partial t} - \operatorname{div}(k_2(u)\nabla(p + p_c(u))) = (1-c)\overline{s} - (1-u)\underline{s}, \tag{5.2}$$

where u is the saturation of the wetting fluid, p its pressure, k_1 and k_2 the reduced mobilities of the two fluids, p_c the capillary pressure, $\overline{s}$ and $\underline{s}$ the injection and extraction velocities and c the injection concentration of the wetting fluid.
Let us define the global pressure p (introduced by Chavent and Jaffré [CJ86]) by

$$\theta := p + \int_0^u \frac{k_2(\tau)}{k_1(\tau)+k_2(\tau)} {p_c}'(\tau)d\tau.$$

Using (3), we easily transform (5.1)-(5.2) in the following equivalent coupled system on $u = u(x,t)$ and $\theta = \theta(x,t)$:

$$\frac{\partial u}{\partial t} - \Delta\varphi(u) + \operatorname{div}(f(u)M(u)\nabla\theta) = c\overline{s} - u\underline{s} \tag{5.3}$$

$$\operatorname{div}(M(u)\nabla\theta) = \overline{s} - \underline{s}, \tag{5.4}$$

where functions φ, f and M are defined by

$$\begin{aligned}
\varphi(x) &:= \int_0^x -\frac{k_1(\tau)k_2(\tau)}{k_1(\tau)+k_2(\tau)}{p_c}'(\tau)d\tau,\\
f(x) &:= \frac{k_1(x)}{k_1(x)+k_2(x)},\\
M(x) &:= k_1(x)+k_2(x).
\end{aligned}$$

We study the coupled system (5.3)-(5.4) on a bounded cylindrical domain $\Omega \times (0,T)$ of $\mathbb{R}^d \times \mathbb{R}^+ (d \geq 1)$, with the following boundary and initial conditions:

$$\nabla\varphi(u)\cdot\mathbf{n} + f(u)M(u)\nabla\theta\cdot\mathbf{n} = 0, \quad x \in \partial\Omega, t \in (0,T), \tag{5.5}$$

$$M(u)\nabla\theta\cdot\mathbf{n} = 0, \quad x \in \partial\Omega, t \in (0,T), \tag{5.6}$$

$$u(\cdot,0) = u_0, \qquad x \in \Omega. \tag{5.7}$$

The function θ is only defined to within an arbitrary function of t by (5.4) so we add the following condition :

$$\int_\Omega \theta = 0, \quad t \in (0,T). \tag{5.8}$$

The aim of this work is to show the convergence of the approximate solution given by the finite volume scheme (5.9)-(5.13) defined in Section 5.2 to a weak solution of (5.3)-(5.4), (5.5)-(5.8) when the size of the mesh and the time step go to zero. We essentially use elliptic and hyperbolic-parabolic finite volume methods developed by R. Eymard, T. Gallouët and R. Herbin. In particular, we admit results proved in [EGH99] and [EGHNS98]. In this study, we assume the following physically reasonable hypotheses on the data:

Hypotheses:

- φ is an increasing Lipschitz continuous function on $[0,1]$, Φ will denote its Lipschitz constant, and $\varphi^* = \max_{x\in[0,1]} |\varphi(x)|$
- f is a nondecreasing Lipschitz continuous function on $[0,1]$, F will denote its Lipschitz constant, $f(0) = 0$ and $f(1) = 1$.
- M is a continuous function on $[0,1]$ with $0 < M_* \leq M(u) \leq M^* < \infty$.
- the functions $\overline{s}$ and $\underline{s}$ belongs to $L^\infty(0,T,L^2(\Omega))$. $\overline{s} \geq 0$ and $\underline{s} \geq 0$ a.e. $x \in \Omega \times (0,T)$, $\int_\Omega \overline{s}(x) - \underline{s}(x)dx = 0$ for a.e. $t \in (0,T)$.
- $u_0 \in L^\infty(\Omega)$, $0 \leq u_0(x) \leq 1$ a.e. $x \in \Omega$.
- c is a constant, $0 \leq c \leq 1$.

Remark 5.1.1 *The function c can also be taken in $L^\infty(Q)$ with $0 \leq c \leq 1$ without any difficulty. One only have to replace c by an approximate in the finite volume scheme as it is done in Chapter 6.*

R.E. Ewing and F. Wheeler [EW84] give *a priori* estimates for Galerkin methods in the case where $\varphi'(u) \geq b_* > 0$. D. Kroener and M. Ohlberger [Ohl97] studied *a posteriori* estimates in the case where $\varphi'(u) = \varepsilon$ and ε is small. In the same direction motivated by local mesh refinements Z.Chen and R.E. Ewing [CE99] recently presented a work on optimal error estimates for degenerate two-phase incompressible flow. Contrary to the works cited below, we do not assume in this paper any existence (nor regularity) of the solutions and use a total finite volume discretization as H.Vignal [Vig96], but without neglecting the capillary pressure.

To obtain discrete estimates, we use techniques which are very similar to the continuous methods. For recent ideas in the theoretical study of this problem, we refer to the G.Gagneux and M.Madaune-Tort's book [GMT96] and Z.Chen and R.E. Ewing's work [CE99]. This problem was also studied with measures as sources terms ([FG00]) which is somewhat more physically admissible than volumetric sources terms.

In Section 2, we present the scheme and we prove a priori estimates on the discrete solution. In Section 3, we use these estimates to obtain compactness properties on the corresponding piecewise constant approximate solution and in Section 4, we prove that the obtained adherence values are weak solutions. In Section 5, we present some numerical results with physically admissible data.

5.2 The finite volume scheme

Assume Ω is a polygonal bounded domain of $\mathbb{R}^d$, $\mathcal{T}$ a mesh of Ω consisting in convex polygonal and δt a time step given by $(N+1)\delta t = T$ where $N \in \mathbb{N}$. The finite volume method consists in integrating the equations over a control volume $K \in \mathcal{T}$ and obtaining a relation between mean values under K and fluxes on the edges of K by using Stokes formula. Thanks to the Neumann boundary conditions, we only need to consider the interfaces between two control volumes. For convection terms, a simple way to get stability is to compute an upwind scheme, but the results also extend to a general monotonous scheme (see [CH99]). For diffusion terms, we have to approximate a normal derivative on the interface. A classical way to solve thsi difficulty is to use a finite element method (see [Ohl97],[CJ86] or [EGHNS98]), but we lose local conservation. Without assumptions on the mesh, we can also take into account the values of u on other control volumes than the two neighbors K and L of the interface (see for example the VF9 method in [Fai92]). The assumptions on the mesh will mainly ensure that the natural and cheap discretization of the flux used in (5.10)-(5.11) is consistent.

5.2.1 Definitions and notations

Definition 5.2.1 (Admissible mesh of Ω) *An admissible mesh $\mathcal{T}$ of Ω is given by a set of open bounded polygonal convex subsets of Ω called control volumes, a family $\mathcal{E}$ of subsets of $\bar{\Omega}$ contained in hyper-planes of $\mathbb{R}^d$ with strictly positive measure, and a family of points (the "centers" of control volumes) satisfying the following properties.*

(i) The closure of the union of all control volumes is $\bar{\Omega}$.

(ii) For any $(K, L) \in \mathcal{T}^2$ with $K \neq L$, either the length of $\bar{K} \cap \bar{L}$ is 0 or $\bar{K} \cap \bar{L} = \bar{\sigma}$ for some $\sigma \in \mathcal{E}$. Then, we will denote $\sigma = K|L$.

(iii) For any $K \in \mathcal{T}$, there exists a subset $\mathcal{E}(K)$ of $\mathcal{E}$ such that $\partial\Omega = \bar{K}\backslash K = \cup_{\sigma\in\mathcal{E}(K)}\bar{\sigma}$. Furthermore, $\mathcal{E} = \cup_{K\in\mathcal{T}}\mathcal{E}(K)$ and we will denote by $\mathcal{N}(K)$ the set of boundary control volumes of K, that is $\mathcal{N}(K) =$

$\{L \in \mathcal{T}, K|L \in \mathcal{E}(K)\}$.

(iv) The family of points $(x_K)_{K\in\mathcal{T}}$ is such that $x_K \in K$ (for all $K \in \mathcal{T}$) and, if $\sigma = K|L$, it is assumed that the straight line (x_K, x_L) is orthogonal to σ.

For a control volume $K \in \mathcal{T}$, we denote by $m(K)$ its measure. If $L \in \mathcal{N}(K)$ we denote by $m(K|L)$ the measure of the interface $K|L$ in $\mathbb{R}^{d-1}$, $d_{K|L}$ the distance between the centers of the control volumes K and L, $T_{K|L} = \frac{m(K|L)}{d_{K|L}}$ the discrete transmissibility and $\mathbf{n}_{K,L}$ the normal vector of $K|L$ outward to K. We denote by $d_{K,K|L}$ the distance between the center x_K of K and the interface $K|L$, and define the size of the mesh by

$$size(\mathcal{T}) = \max_{K\in\mathcal{T}} diam(K)$$

To prove convergence Theorem 5.3.2, we need uniform regularity properties on meshes in the following sense:

Definition 5.2.2 *An admissible mesh $\mathcal{T}$ is ξ-regular if for all $K \in \mathcal{T}$,*

$$\sum_{L\in\mathcal{N}(K)} m(K|L)d_{K|L} \leq m(K)\,\xi.$$

5.2.2 The scheme

Let $\mathcal{T}$ be an admissible mesh and δt a time step such that $T = (N+1)\delta t$ with $N \in \mathbb{N}$. We define $\overline{S}_K^{n+1}$ and $\underline{S}_K^{n+1}$ by

$$\overline{S}_K^{n+1} = \frac{1}{\delta t}\int_{n\delta t}^{(n+1)\delta t}\int_K \overline{s},$$
$$\underline{S}_K^{n+1} = \frac{1}{\delta t}\int_{n\delta t}^{(n+1)\delta t}\int_K \underline{s}.$$

With the notations previously introduced, one may define a finite volume scheme as the following set of equations for the discrete unknowns (U, Θ) where $U = (U_K^n)_{K\in\mathcal{T}, n\in[\![0,N+1]\!]}$ and $\Theta = (\Theta_K^n)_{K\in\mathcal{T}, n\in[\![1,N+1]\!]}$:

$\forall K \in \mathcal{T}$,

$$U_K^0 = \frac{1}{m(K)}\int_K u_0(x)dx. \tag{5.9}$$

$\forall K \in \mathcal{T}, \forall n \in [\![0, N]\!]$,

$$\frac{U_K^{n+1} - U_K^n}{\delta t} m(K) - \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1})) + \sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^{n+1} f(u)_{K|L}^{n+1} = c\overline{S}_K^{n+1} - U_K^{n+1}\,\underline{S}_K^{n+1} \tag{5.10}$$

$\forall K \in \mathcal{T}, \forall L \in \mathcal{N}(K), \forall n \in [\![0, N]\!]$,

$$\mathbf{q}_{K,L}^{n+1} = -M(u)_{K|L}^{n+1} T_{K|L}(\Theta_L^{n+1} - \Theta_K^{n+1}), \tag{5.11}$$

$\forall K \in \mathcal{T}, \forall n \in [\![0, N]\!]$,

$$\sum_{L \in \mathcal{N}(K)} \mathbf{q}_{K,L}^{n+1} = \overline{S}_K^{n+1} - \underline{S}_K^{n+1}, \tag{5.12}$$

$\forall n \in [\![0, N]\!]$,

$$\sum_{K \in \mathcal{T}} m(K) \Theta_K^{n+1} = 0, \tag{5.13}$$

where $f(u)_{K|L}^{n+1}$ and $M(u)_{K|L}^{n+1}$ are respectively an upwind discretization of $f(u)$ and a consistent approximation of $M(u)$ on the interface $K|L$ given by

$$f(u)_{K|L}^{n+1} = \begin{cases} f(U_K^{n+1}) & \text{if} \quad \mathbf{q}_{K,L}^{n+1} > 0 \\ \\ f(U_L^{n+1}) & \text{if} \quad \mathbf{q}_{K,L}^{n+1} < 0 \end{cases},$$

$$M(u)_{K|L}^{n+1} = \frac{d_{K|L}}{\frac{d_{K,K|L}}{M(U_K^{n+1})} + \frac{d_{L,K|L}}{M(U_L^{n+1})}}. \tag{5.14}$$

Remark 5.2.1 *Definition (5.14) of $M(u)_{K|L}^{n+1}$ by a harmonic mean ensures consistency property in the general case when the function $M(u)$ is discontinuous on the interface $K|L$ (see [Her96],[EGH00b]). However, in our proof of convergence we only need $M(u)_{K|L}^{n+1}$ to be in the interval $[M(U_K^{n+1}), M(U_L^{n+1})]$ since the functions used are more regular.*

5.2.3 *A priori* estimates

The scheme (5.9)-(5.13) is time implicit so the existence of a solution must be proven. We will first prove *a priori* estimates assuming existence of a solution and then prove the existence by using the Leray Schauder Theorem. We also use these estimates to obtain compactness properties.

Proposition 5.2.1 *Assume that $((U_K^n)_{K \in \mathcal{T}, n \in [\![0,N+1]\!]}, (\Theta_K^{n+1})_{K \in \mathcal{T}, n \in [\![0,N]\!]})$ is a solution to (5.9)-(5.13) then*

$$0 \le U_K^{n+1} \le 1, \qquad \forall K \in \mathcal{T}, \forall n \in [\![0, N+1]\!]. \tag{5.15}$$

Moreover there exist $C_1(u_0, \overline{s}, \underline{s}, \Phi) \ge 0$ and $C_2(M_*, \overline{s}, \underline{s}) > 0$ such that

$$\sum_{n=0}^{N} \delta t \frac{1}{2} \sum_{K \in \mathcal{T}} \sum_{L \in \mathcal{N}(K)} T_{K|L} (\varphi(U_L^{n+1}) - \varphi(U_K^{n+1}))^2 \le C_1 \tag{5.16}$$

and $$\sum_{K \in \mathcal{T}} \sum_{L \in \mathcal{N}(K)} T_{K|L} (\Theta_L^{n+1} - \Theta_K^{n+1})^2 \le C_2, \forall n \in [\![0, N]\!]. \tag{5.17}$$

Proof. Rewriting the discrete convection flux in a non divergence form using (5.14) and (5.12), we obtain

$\forall K \in \mathcal{T}, \forall n \in [\![0, N]\!]$,

$$\sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^{n+1} f(u)_{K|L}^{n+1} = -\sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^{n+1^-}(f(U_L^{n+1}) - f(U_K^{n+1})) + f(U_K^{n+1})(\overline{S}_K^{n+1} - \underline{S}_K^{n+1}), \tag{5.18}$$

where one denotes by $x^- = \max(0, -x)$.

In order to prove the discrete maximum principle (5.15), we follow the continuous case. If U attains its bounds on $\Omega \times \{0\}$, *i.e.* at points of type $(K, 0)$, definition (5.9) of U_K^0 gives the conclusion. By contradiction, if for example $max(U) > U_S$ then necessarily U attains its maximum at an interior point of the parabolic domain $Q = \Omega \times [0, T)$, *i.e.* of type $(K, n+1)$. In that case, by (5.18) and (5.10) we have

$$\begin{aligned} \frac{U_K^{n+1} - U_K^n}{\delta t} m(K) \;&+ \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1}) - \varphi(U_L^{n+1})) \\ &+ \sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^{n+1^-}(f(U_K^{n+1}) - f(U_L^{n+1})) \\ &+ (f(U_K^{n+1}) - c)\overline{S}_K^{n+1} + (U_K^{n+1} - f(U_K^{n+1}))\,\underline{S}_K^{n+1} = 0, \end{aligned} \tag{5.19}$$

and φ and f are nondecreasing functions, so we have $U_K^n \geq U_K^{n+1}$. Consequently, also U attain its maximum at point (K, n) and by induction, the maximum is attained on $\Omega \times \{0\}$ which leads to a contradiction.

Proofs of the discrete energy estimates (5.16) and (5.17) also mimic continuous ones. Multiplying (5.19) by $\delta t\, U_K^{n+1}$ and summing the result over $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$ yields $E1 + E2 + E3 + E4 = 0$ with

$$\begin{aligned} E1 &= \sum_{n=0}^{N} \sum_{K\in\mathcal{T}} m(K)(U_K^{n+1} - U_K^n)U_K^{n+1}, \\ E2 &= \sum_{n=0}^{N} \delta t \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1}) - \varphi(U_L^{n+1}))U_K^{n+1}, \\ E3 &= \sum_{n=0}^{N} \delta t \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \mathbf{q}_{K,L}^{n+1^-}(f(U_K^{n+1}) - f(U_L^{n+1}))U_K^{n+1}, \\ E4 &= \sum_{n=0}^{N} \delta t \sum_{K\in\mathcal{T}} (f(U_K^{n+1}) - c)U_K^{n+1}\overline{S}_K^{n+1} + (U_K^{n+1} - f(U_K^{n+1}))U_K^{n+1}\,\underline{S}_K^{n+1}. \end{aligned}$$

By a discrete time integration by parts, we obtain

$$\begin{aligned} E1 &= \frac{1}{2}\sum_{K\in\mathcal{T}} m(K)(u_K^{N+1})^2 - \frac{1}{2}\sum_{K\in\mathcal{T}} m(K)(u_K^0)^2 + \frac{1}{2}\sum_{n=0}^{N}\sum_{K\in\mathcal{T}} m(K)(U_K^{n+1} - U_K^n)^2 \\ &\geq -\frac{1}{2}\|u_0\|_{L^2(\Omega\times(0,T))}^2. \end{aligned}$$

Gathering by edges we get for $E2$

$$\begin{aligned} E2 &= \sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1})-\varphi(U_L^{n+1}))(U_K^{n+1}-U_L^{n+1}) \\ &\geq \frac{1}{\Phi}\sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1})-\varphi(U_L^{n+1}))^2. \end{aligned}$$

To deal with $E3$, we need a technical lemma which is proved for example in [EGH00b].

Lemma 5.2.1 *Let f be a nondecreasing continuous function on $\mathbb{R}$ and define g by $g(u) = uf(u) - \int_0^u f(\tau)d\tau$. Then for every $(a,b)\in\mathbb{R}^2$,*

$$(f(a)-f(b))a \geq g(a)-g(b)$$

By using Lemma 5.2.1 and the local conservation property $\mathbf{q}_{K,L}^{n+1}+\mathbf{q}_{L,K}^{n+1}=0$ we get

$$\begin{aligned} E3 &\geq \sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}} \mathbf{q}_{K,L}^{n+1^-}(g(U_L^{n+1})-g(U_K^{n+1})) \\ &= \sum_{K\in\mathcal{T}} g(U_K^{n+1})\sum_{L\in\mathcal{N}(K)}\mathbf{q}_{K,L}^{n+1} \\ &= -\sum_{K\in\mathcal{T}} g(U_K^{n+1})(\overline{S}_K^{n+1}-\underline{S}_K^{n+1}). \end{aligned}$$

Hence:

$$\begin{aligned} E3+E4 &\geq \sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}}\overline{S}_K^{n+1}(f(U_K^{n+1})U_K^{n+1}-g(U_K^{n+1})-cU_K^{n+1}) \\ &\quad +\sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}}\underline{S}_K^{n+1}(g(U_K^{n+1})-f(U_K^{n+1})U_K^{n+1}+(U_K^{n+1})^2) \\ &\geq \sum_{n=0}^{N}\delta t\sum_{K\in\mathcal{T}} -2(\overline{S}_K^{n+1}+\underline{S}_K^{n+1}) \\ &\geq -2(\|\overline{s}\|_{L^1(\Omega\times(0,T))}+\|\underline{s}\|_{L^1(\Omega\times(0,T))}). \end{aligned}$$

Collecting the previous inequalities yields exactly (5.16) with

$$C_1 = \Phi(\frac{1}{2}\|u_0\|^2_{L^2(\Omega\times(0,T))}+2(\|\overline{s}\|_{L^1(\Omega\times(0,T))}+\|\underline{s}\|_{L^1(\Omega\times(0,T))})).$$

Now let us multiply (5.12) by Θ_K^{n+1} and sum over $K\in\mathcal{T}$. Gathering by edges, we obtain

$$\begin{aligned} \sum_{K\in\mathcal{T}}\frac{1}{2}\sum_{L\in\mathcal{N}(K)} M(u)_{K|L}^{n+1}(\Theta_K^{n+1}-\Theta_L^{n+1})^2 &= \sum_{K\in\mathcal{T}}(\overline{S}_K^{n+1}-\underline{S}_K^{n+1})\Theta_K^{n+1} \\ &\leq \|\overline{s}-\underline{s}\|_{L^\infty(0,T,L^2(\Omega))}(\sum_{K\in\mathcal{T}} m(K)(\Theta_K^{n+1})^2)^{\frac{1}{2}} \end{aligned}$$

And by the discrete Poincaré inequality (see [CVV99]), there exists $C(\Omega)$ such that

$$\sum_{K\in\mathcal{T}} m(K)(\Theta_K^{n+1})^2 \leq C(\Omega)^2 \sum_{K\in\mathcal{T}} \frac{1}{2} \sum_{L\in\mathcal{N}(K)} T_{K|L}(\Theta_K^{n+1} - \Theta_L^{n+1})^2.$$

This gives (5.17) with $C_2 = C(\Omega)\frac{1}{M_*}\|\overline{s} - \underline{s}\|^2_{L^2(\Omega\times(0,T))}$.
□

5.2.4 Existence of a solution to the scheme

Let $E = \mathbb{R}^{[\![0,N+1]\!]\times\mathcal{T}} \times \mathbb{R}^{[\![1,N+1]\!]\times\mathcal{T}}$ and $f : E \to E$ that associate to (U,Θ) the solution $(\tilde{U},\tilde{\Theta})$ of the following set of equations

$\forall K \in \mathcal{T}$,

$$\tilde{u}_K^0 = \frac{1}{m(K)} \times \int_K u_0(x)dx,$$

$\forall K \in \mathcal{T}$ et $\forall n \in [\![0,N]\!]$

$$\begin{aligned}\frac{\tilde{U}_K^{n+1} - \tilde{U}_K^n}{\delta t} m(K) \quad &- \sum_{\sigma\in\mathcal{E}(K)} T_{K|L}(\varphi(U_L^{n+1}) - \varphi(U_K^{n+1})) \\ &- \sum_{\sigma\in\mathcal{E}(K)} \tilde{\mathbf{q}}_{K,L}^{n+1+} f(U_K^{n+1}) - \tilde{\mathbf{q}}_{K,L}^{n+1-} f(U_L^{n+1}) = c\overline{S}_K^{n+1} - U_K^{n+1}\,\underline{S}_K^{n+1},\end{aligned}$$

$\forall K \in \mathcal{T}$, $\forall L \in \mathcal{N}(K)$, $\forall n \in [\![0,N]\!]$

$$\mathbf{q}_{K,L}^{n+1} = -M(u)_{K|L}^{n+1} T_{K|L}(\Theta_L^{n+1} - \Theta_K^{n+1}),$$

$$\tilde{\mathbf{q}}_{K,L}^{n+1} = -M(u)_{K|L}^{n+1} T_{K|L}(\tilde{\Theta}_L^{n+1} - \tilde{\Theta}_K^{n+1}),$$

$\forall K \in \mathcal{T}, \forall n \in [\![0,N]\!]$,

$$\sum_{L\in\mathcal{N}(K)} \tilde{\mathbf{q}}_{K,L}^{n+1} = \overline{S}_K^{n+1} - \underline{S}_K^{n+1},$$

$\forall n \in [\![0,N]\!]$,

$$\sum_{K\in\mathcal{T}} m(K)\tilde{\Theta}_K^{n+1} = 0.$$

Since $M(u)_{K|L}^{n+1} \geq M* > 0$, the linear system on $\tilde{\Theta}$ is strictly elliptic uniformly with respect to (U,Θ), so that f is well defined on E and is continuous. Moreover for any $t \in [0,1]$, by construction, each solution of $(U,\Theta) = tf(U,\Theta)$ is a solution of (5.9)-(5.13) with $t\varphi$, tu_0, $t\overline{s}$ and $t\underline{s}$ instead of φ, u_0, $\overline{s}$ and $\underline{s}$. Thus it also satisfies estimates obtained at Proposition 5.2.1 if $t \in [0,1]$ and from the Leray Schauder fixed point Theorem (5.9)-(5.13) has at least one solution.

5.3 Convergence results

To each solution $(U,\Theta)_{\mathcal{T},\delta t}$ of (5.9)-(5.13) for an admissible mesh $\mathcal{T}$ and a time step δt corresponds an approximate solution $(u_{\mathcal{T},\delta t},\theta_{\mathcal{T},\delta t})$ of problem (5.3)-(5.4), (5.5)-(5.8) defined a.e. on $\Omega\times(0,T)$ by

$$\begin{aligned} u_{\mathcal{T},\delta t}(x,t) &= U_K^{n+1}, \qquad x\in K, t\in(n\delta t,(n+1)\delta t),\\ \theta_{\mathcal{T},\delta t}(x,t) &= \Theta_K^{n+1}, \qquad x\in K, t\in(n\delta t,(n+1)\delta t). \end{aligned}$$

The first step in direction to the convergence Theorem consists in the proof of compactness properties on $u_{\mathcal{T},\delta t}$ and $\theta_{\mathcal{T},\delta t}$, by using *a priori* estimates on the discrete solution obtained in Proposition 5.2.1.

5.3.1 Compactness of $u_{\mathcal{T},\delta t}$

We shall prove that $\varphi(u_{\mathcal{T},\delta t})$ is relatively compact in $L^2(\Omega\times(0,T))$ for the strong topology by using Kolmogorov's Theorem and that when $size(\mathcal{T})\to 0$ and $\delta t\to 0$, the limit of each convergent sequence of approximate solutions belongs to $L^2(0,T,H^1(\Omega))$. Let us first recall Kolmogorov's compactness Theorem, which is a consequence of the Ascoli compactness Theorem.

Theorem 5.3.1 (Fréchet-Kolmogorov) *Let $\mathcal{F}$ be a bounded subspace of $L^2(\mathbb{R}^d)$ and Ω a bounded domain of $\mathbb{R}^d$, then $\mathcal{F}$ is relatively compact in $L^2(\Omega)$ if and only if*

$$\lim_{|\xi|\to 0}\sup_{f\in\mathcal{F}}\|f(\cdot+\xi)-f(\cdot)\|_{L^2(\mathbb{R}^d)}=0.$$

In our case, to apply this Theorem on $Q=\Omega\times(0,T)$, we need to study the space and time translates of $\varphi(u_{\mathcal{T},\delta t})$. As a direct consequence of (5.16) (see [EGH99],[EGH00b]) we already have the following result.

Proposition 5.3.1 (Space translates) *Let C_1 be defined in Proposition 5.2.1, then $\forall\xi\in\mathbb{R}^d$,*

$$\int_0^T\int_{\Omega_\xi}[\varphi(u_{\mathcal{T},\delta t}(x+\xi,\cdot)-\varphi(u_{\mathcal{T},\delta t}(x,\cdot))]^2dx\le C\Phi|\xi|(2m(\mathcal{T})+|\xi|),$$

where $\Omega_\xi=\{x\in\Omega,[x,x+\xi]\subset\Omega\}$, and $|\xi|$ the Euclidean norm on $\mathbb{R}^d$.

We can establish an analog but slightly different result for time translates estimates. We now adapt the method of [EGHNS98], but since this method is not quite well known, we shall give the complete proof of it. Let us first state exactly the result we shall prove.

Proposition 5.3.2 (Time translates) *There exist $C'(\varepsilon,\varphi,f,\mathbf{q},u_0,\Omega,T)>0$ such that for every $s\in\mathbb{R}^+$,*

$$\int_0^{T-s}\int_\Omega(\varphi(u_{\mathcal{T},\delta t}(x,t+s))-\varphi(u_{\mathcal{T},\delta t}(x,t))\,)^2dxdt\le C's.$$

Proof. Let us define $A(t)=\int_\Omega(u_{\mathcal{T},\delta t}(x,t+s)-u_{\mathcal{T},\delta t}(x,t))(\varphi(u_{\mathcal{T},\delta t}(x,t+s)-\varphi(u_{\mathcal{T},\delta t}(x,t)))dxdt$. Then

$$\int_\Omega(\varphi(u_{\mathcal{T},\delta t}(x,t+s))-\varphi(u_{\mathcal{T},\delta t}(x,t))\,)^2dxdt\le A(t)\,\Phi.$$

If for any $t \in \mathbb{R}$ we denote by $n(t)$ the integer part of $\frac{t}{\delta t}$, then for any $K \in \mathcal{T}$ and $x \in K$,

$$\begin{aligned} u_{\mathcal{T},\delta t}(x,t+s) - u_{\mathcal{T},\delta t}(x,t) &= \sum_{n=n(t)}^{n(t+s)-1} U_K^{n+1} - U_K^n \\ &= \frac{1}{m(K)} \sum_{n=n(t)}^{n(t+s)-1} \delta t \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1}) - \varphi(U_L^{n+1})) - \mathbf{q}_{K,L}^{n+1} f(u)_{K|L}^{n+1}. \end{aligned}$$

So $A(t) = A_1(s,t) - A_1(0,t) - A_2(s,t) + A_2(0,t)$ where for $\rho = 0$ or $\rho = s$, we denote by

$$\begin{aligned} A_1(\rho,t) &= \sum_{n=n(t)}^{n(t+s)-1} \delta t \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1}) - \varphi(U_L^{n+1}))\varphi(U_K^{n(t+\rho)}), \\ A_2(\rho,t) &= \sum_{n=n(t)}^{n(t+s)-1} \delta t \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}(K)} \mathbf{q}_{K,L}^{n+1} f(u)_{K|L}^{n+1}\varphi(U_K^{n(t+\rho)}). \end{aligned}$$

Gathering by edges and using the local conservation of the discrete fluxes, we get

$$\begin{aligned} A_1(\rho,t) &= \sum_{n=n(t)}^{n(t+s)-1} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1}) - \varphi(U_L^{n+1}))(\varphi(U_K^{n(t+\rho)}) - \varphi(U_L^{n(t+\rho)})), \\ A_2(\rho,t) &= \sum_{n=n(t)}^{n(t+s)-1} \delta t \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}(K)} \mathbf{q}_{K,L}^{n+1} f(u)_{K|L}^{n+1}(\varphi(U_K^{n(t+\rho)}) - \varphi(U_L^{n(t+\rho)})). \end{aligned}$$

Now by using Young inequality,

$$\begin{aligned} |A_1(\rho,t)| &\leq \frac{1}{2} \sum_{n=n(t)}^{n(t+s)-1} \delta t\,(S(n) + S(n(t+\rho))), \\ |A_2(\rho,t)| &\leq \frac{1}{2} \sum_{n=n(t)}^{n(t+s)-1} \delta t\,(R(n) + S(n(t+\rho))), \end{aligned}$$

where

$$\begin{aligned} S(n) &= \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} T_{K|L}(\varphi(U_K^{n+1}) - \varphi(U_L^{n+1}))^2, \\ R(n) &= \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \frac{1}{T_{K|L}}(\mathbf{q}_{K,L}^{n+1} f(u)_{K|L}^{n+1})^2. \end{aligned}$$

Let us assume the following technical results.

Lemma 5.3.1 *Let $B : \mathbb{N} \to \mathbb{R}^+$ such that $\sum_{n=0}^{N} \delta t B(n) \leq C$, then $\forall s \in \mathbb{R}^+$, $\int_0^{T-s} \sum_{n=n(t)}^{n(t+s)-1} \delta t B(n) dt \leq C\, s$.*

Lemma 5.3.2 *Let $B : \mathbb{N} \to \mathbb{R}^+$ such that $\displaystyle\sum_{n=0}^{N} \delta t B(n) \leq C$, then $\forall s \in \mathbb{R}^+$, $\displaystyle\int_0^{T-s} \sum_{n=n(t)}^{n(t+s)-1} \delta t B(n(t+\rho))dt \leq C\, s$.*

According to estimates (5.16) and (5.17) S and R satisfies the assumptions of the two lemmas, so we obtain

$$\begin{aligned}\|A_1(\rho,t)\| &\leq C',\\ \|A_2(\rho,t)\| &\leq C'',\end{aligned}$$

and collecting the previous inequalities complete the proof of Proposition 5.3.2. Yet, it only remains to show the two lemmas.
For Lemma 5.3.1, $\displaystyle\int_0^{T-s} \sum_{n=n(t)}^{n(t+s)-1} \delta t B(n) = \sum_{n=0}^{N} \delta t B(n) \int_0^{T-s} \mathbb{1}_{n \in [\![n(t),n(t+s)-1]\!]} dt$ and $n \in [\![n(t), n(t+s)-1]\!]$ if and only if $t \in [(n+1)\delta t - s, (n+1)\delta t)$ so $\displaystyle\int_0^{T-s} \mathbb{1}_{n \in [\![n(t),n(t+s)-1]\!]} dt \leq s$.

For Lemma 5.3.2, $\displaystyle\int_0^{T-s} \sum_{n=n(t)}^{n(t+s)-1} \delta t B(n(t+\rho))dt = \sum_{n=0}^{N} \delta t B(n) \int_0^{T-s} (n(t+s) - n(t))\, \mathbb{1}_{n(t+\rho)=n} dt$, but $n(t+s)-n(t)$ is periodic with period δt so $\int_0^{T-s} (n(t+s)-n(t))\, \mathbb{1}_{n(t+\rho)=n} dt \leq \int_{n\delta t-\rho}^{(n+1)\delta t-\rho} (n(t+s)-n(t))dt \leq \int_0^{\delta t} n(t+s)dt \leq s$.
□

To prove compactness of $\varphi(u_{\mathcal{T},\delta t})$ in $L^2(\Omega)$, let us check hypothesis of Theorem 5.3.1. From Propositions 5.3.1 and 5.3.2 we easily deduce that if we extend $u_{\mathcal{T},\delta t}$ by zero outside of $\Omega \times (0,T)$, for every $\xi \in \mathbb{R}^d$ and $s \in \mathbb{R}^+$ one has

$$\begin{aligned}\|\varphi(u_{\mathcal{T},\delta t}(\cdot+\xi,\cdot+s)) - \varphi(u_{\mathcal{T},\delta t}(\cdot,\cdot))\|_{L^2(\mathbb{R}^{m+1})} \leq\ & 2C|\xi|(|\xi|+2h) + 2C's \\ & +(4T|\xi|m(\partial\Omega) + 2m(\Omega)s)(\varphi^*)^2.\end{aligned}$$

Lets $u_m = u_{\mathcal{T}_m,\delta t_m}$ be a sequence of approximate solutions with $size(\mathcal{T}_m) \to 0$ when m tends to infinity and suppose that $\varphi(u_m)$ tends to $\bar{\varphi}$ in $L^2(\Omega \times (0,T))$. It remains to show that $\bar{\varphi}$ is in $L^2(0,T,H^1(\Omega))$. In order to prove it, we shall use space translate estimates in the interior of Ω.
Let $\omega \subset\subset \Omega$. From Proposition 5.3.1, if $(\xi, z) \in \mathbb{R}^d \times \mathbb{R}$ satisfies $|\xi| \leq d(\omega, \mathbb{R}^d - \Omega)$ and $-1 \leq z \leq 1$,

$$\left\| \frac{\varphi(u_m(\cdot+z\xi,\cdot)) - \varphi(u_m(\cdot,\cdot))}{z} \right\|_{L^2(\omega\times(0,T))} \leq |\xi|\sqrt{C\Phi} + \sqrt{2h_m \frac{|\xi|}{|z|}}.$$

so by letting m tends to infinity, it holds

$$\left\| \frac{\bar{\varphi}(\cdot+z\xi,\cdot) - \bar{\varphi}(\cdot,\cdot)}{z} \right\|_{L^2(\omega\times(0,T))} \leq |\xi|\sqrt{C\Phi},$$

and by letting z tends to zero we obtain finally

$$\|\nabla\bar{\varphi}\cdot\xi\|_{L^2(\omega\times(0,T))} \leq |\xi|\sqrt{C\Phi}. \tag{5.20}$$

By homogeneity, inequality (5.20) is true for every $\xi \in \mathbb{R}^d$ so $\bar{\varphi} \in L^2(0,T,H^1(\Omega))$ and $\|\nabla\bar{\varphi}\|_{L^2((0,T)\times\Omega)} \leq \sqrt{C\Phi}$. This regularity property will be useful for example to make a sense to a weak formulation for problem (5.3)-(5.4), (5.5)-(5.8) .

5.3.2 Compactness of $\theta_{\mathcal{T},\delta t}$

In the same way as for $\varphi(u_{\mathcal{T},\delta t})$, we can show space translate estimates on $\theta_{\mathcal{T},\delta t}$, but since the equation satisfied by θ does not include time derivatives relative to θ, we do not have any time translate estimate. Hence we cannot apply the same method to obtain compactness on $\theta_{\mathcal{T},\delta t}$. However, by Poincaré inequality, $(\theta_{\mathcal{T},\delta t})_{\mathcal{T},\delta t}$ is a bounded in $L^\infty(0,T,L^2(\Omega))$. Therefore $\theta_{\mathcal{T},\delta t}$ is sequentially weakly relatively compact in $L^2(\Omega\times(0,T))$ and by the same arguments as previously, every possible limit when $size(\mathcal{T})$ tends to zero belongs to $L^\infty(0,T,H^1(\Omega))$. This is sufficient for convergence under the hypothesis that φ' is strictly nondecreasing, since we get strong convergence of $u_{\mathcal{T},\delta t}$.

5.3.3 Convergence theorem

Definition 5.3.1 (Weak solution) *(u,θ) is a weak solution of Problem (5.3)-(5.4), (5.5)-(5.8) if $u \in L^\infty(\Omega\times(0,T))$, $0 \leq u(x,t) \leq 1$ a.e $(x,t) \in \Omega\times(0,T)$, $\varphi(u) \in L^2(0,T,H^1(\Omega))$, $\theta \in L^\infty(0,T,H^1(\Omega))$ and for any $\psi \in (\mathbb{R}^d \times [0,T))^2$, we have:*

$$\int_0^T\!\!\int_\Omega u\psi_t - \int_0^T\!\!\int_\Omega \nabla\varphi(u)\cdot\nabla\psi - \int_0^T\!\!\int_\Omega f(u)M(u)\nabla\theta\cdot\nabla\psi + \int_\Omega u_0\psi(\cdot,0) = -\int_0^T\!\!\int_\Omega c\overline{s}\psi + \int_0^T\!\!\int_\Omega u\underline{s}\psi \quad (5.21)$$

$$\int_0^T\!\!\int_\Omega M(u)\nabla\theta\cdot\nabla\psi = \int_0^T\!\!\int_\Omega (\overline{s}-\underline{s})\psi \quad (5.22)$$

We shall now give and prove the main result of this paper.

Theorem 5.3.2 (The convergence theorem) *Let $(u_m,\theta_m) = (u_{\mathcal{T}_m,\delta t_m}, \theta_{\mathcal{T}_m,\delta t_m})$ be a sequence of approximate solutions given by scheme (5.9)-(5.13) . Let us assume that there exists $\xi > 0$ such that for every $m \in \mathbb{N}$, $\mathcal{T}_m$ is a ξ-regular admissible mesh. Assume also $size(\mathcal{T}_m) \to 0$ and $\delta t_m \to 0$ when m tends to zero. Then there exists a weak solution (u,θ) of Problem (5.3)-(5.4), (5.5)-(5.8) , such that up to a subsequence,*

$$u_m \to u, \quad \text{strongly in } L^2(\Omega\times(0,T)) \text{ as } m\to\infty,$$
$$\theta_m \to \theta, \quad \text{weakly in } L^2(\Omega\times(0,T)) \text{ as } m\to\infty.$$

Remark 5.3.1 *Under the assumption that $f(\varphi^{-1})$ is Holder continuous with exponent $\frac{1}{2}$ and with the additional hypothesis $\|\nabla\theta\| \in L^\infty(\Omega\times(0,T))$, we can prove that the weak solution is unique (see [CE99]). So the whole sequence of approximate solutions is convergent.*

Proof. From compactness properties of u_m and θ_m, we already know that up to a subsequence $\varphi(u_m) \to \bar{\varphi}$ strongly in $L^2(\Omega\times(0,T))$, $\theta_m \to p$ weakly in $L^2(\Omega\times(0,T))$ as $m\to\infty$. Since $size(\mathcal{T}_m) \to 0$, $\bar{\varphi} \in L^2(0,T,H^1(\Omega))$ and $\theta \in L^\infty(0,T,H^1(\Omega))$. Now since φ is strictly nondecreasing, by using for example the dominated convergence Theorem we deduce that u_m tends to $u = \varphi^{-1}(\bar{\varphi})$ strongly in $L^2(\Omega\times(0,T))$. It remains to show that u is a weak solution of Problem (5.3)-(5.4), (5.5)-(5.8) .

Let $(\mathcal{T},\delta t)=(\mathcal{T}_m,\delta t_m)$. We define the discretization and approximate of ψ denoted respectively Ψ and $\psi_{\mathcal{T},\delta t}$ by the following formulas:

$$\Psi_K^{n+1}=\psi(x_K,n\delta t),\qquad K\in\mathcal{T},n\in[\![0,N]\!], \tag{5.23}$$

$$\psi_{\mathcal{T}}(x,t)=\Psi_K^{n+1},\qquad x\in K,t\in(n\delta t,(n+1)\delta t). \tag{5.24}$$

In order to prove that (u,θ) is a weak solution, we multiply (5.10) and (5.12) by $\delta t\Psi_K^{n+1}$ and sum over $n\in[\![0,N]\!]$ and $K\in\mathcal{T}$. Then we let m tend to infinity and show that we obtain (5.21) and (5.22) when passing to the limit.
As in [EGHNS98],[Vig96], [CH99], by using the consistency of fluxes, we obtain at the limit the following terms.

$$\int_0^T\!\!\int_\Omega u\psi_t+\int_\Omega u_0,\int_0^T\!\!\int_\Omega\nabla\varphi(u)\cdot\nabla\psi,\int_0^T\!\!\int_\Omega c\overline{s}\psi,\int_0^T\!\!\int_\Omega u\underline{s}\psi\quad\text{and}\quad\int_0^T\!\!\int_\Omega(\overline{s}-\underline{s})\psi.$$

But we encounter original difficulties to obtain the two last terms, namely:

$$\int_0^T\!\!\int_\Omega f(u)M(u)\nabla\theta\ \ and\ \int_0^T\!\!\int_\Omega M(u)\nabla\theta.$$

Indeed, heuristically, we have to prove the weak convergence of $f(u_m)M(u_m)\nabla\theta_m$ and $M(u_m)\nabla\theta_m$ to $f(u)M(u)\nabla\theta$ and $M(u)\nabla\theta$, with u_m strongly convergent to u and θ_m bounded in $L^2(0,T,H^1(\Omega))$ and weakly convergent to θ. In the continuous case, this problem can be solved since a product of two functions, the first converging strongly and the other weakly, is weakly convergent to the product of the limits. However the gradient of discrete function is not in general a function, so we need to use a regularization argument that we shall detail.

Remark 5.3.2 *Our heuristic argument justifies why the strong convergence of u_m is crucial. Indeed, a product of weak convergent function in general does not converge to the product of the limit, even if it has a weak limit. In our case we obtain this strong convergence by using the hypothesis that φ' is a strictly increasing function, but our method would also work in all cases where we were able to prove that $f(u_m)$ and $M(u_m)$ converge strongly.*

We will restrict ourselves to the proof concerning $A=\int_\Omega\int_0^T f(u)M(u)\nabla\theta\cdot\nabla\psi$ because the other integral is a particular case with $f=1$. Let us first define $A_{\mathcal{T},\delta t}$ by

$$A_{\mathcal{T},\delta t}=\sum_{n=0}^N\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}f(u)_{K|L}^{n+1}M(u)_{K|L}^{n+1}(\Theta_L^{n+1}-\Theta_K^{n+1})(\Psi_L^{n+1}-\Psi_K^{n+1}).$$

$A_{\mathcal{T},\delta t}$ is the term corresponding to A when we multiply (5.10) by $\delta t\Psi_K^{n+1}$, sum over $K\in\mathcal{T}$ and $n\in[\![0,N]\!]$, and gather by edges. Assume first that $f(u)$ and $M(u)$ are in $\mathcal{D}(\Omega\times(0,T))$, and $M(U)$, $m(u_{\mathcal{T},\delta t})$ and $f(U)$, $f(u_{\mathcal{T},\delta t})$ are discretizations and approximations of $f(u)$ and $M(u)$ defined in the same way as Ψ and $\psi_{\mathcal{T},\delta t}$ by (5.23)-(5.24). By the Stokes formula,

$$A=-\int_0^T\!\!\int_\Omega\theta\ \mathrm{div}(f(u)M(u)\nabla\psi).$$

So because of the weak convergence of θ_m to θ, $A=\lim_{m\to\infty}B_{\mathcal{T}_m,\delta t_m}$ where $B_{\mathcal{T},\delta t}$ is given by

$$B_{\mathcal{T},\delta t}=\int_0^T\!\!\int_\Omega\theta_{\mathcal{T},\delta t}\mathrm{div}(f(u)M(u)\nabla\psi),$$

and by definition of the piecewise constant function $\theta_{\mathcal{T},\delta t}$, we get

$$
\begin{aligned}
B_{\mathcal{T},\delta t} &= -\sum_{n=0}^{N}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}\Theta_K^{n+1}\int_{n\delta t}^{(n+1)\delta t}\int_{K|L} f(u)M(u)\nabla\psi\cdot\mathbf{n}_{K,L}\\
&= \sum_{n=0}^{N}\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}(\Theta_L^{n+1}-\Theta_K^{n+1})\int_{n\delta t}^{(n+1)\delta t}\int_{K|L} f(u)M(u)\nabla\psi\cdot\mathbf{n}_{K,L}.
\end{aligned}
$$

Now, let us compare $B_{\mathcal{T},\delta t}$ and $A_{\mathcal{T},\delta t}$, using the consistency of the flux on the interfaces $K|L$ for regular functions. By the Cauchy-Schwarz inequality, we have

$$
|A_{\mathcal{T},\delta t}-B_{\mathcal{T},\delta t}|^2 \le \sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}T_{K|L}(\Theta_K^{n+1}-\Theta_L^{n+1})^2\sum_{n=0}^{N}\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}\frac{1}{T_{K|L}}R_{K|L}^2,
$$

where

$$
R_{K|L} = |f(u)_{K|L}^{n+1}M(u)_{K|L}^{n+1}T_{K|L}(\Psi_K^{n+1}-\Psi_L^{n+1}) - \frac{1}{\delta t}\int_{n\delta t}^{(n+1)\delta t}\int_{K|L} f(u)M(u)\nabla\psi\cdot\mathbf{n}_{K,L}|.
$$

By using the regularity of $f(u)$, $M(u)$ and ψ and the orthogonality property $x_K - x_L = d_{K|L}\mathbf{n}_{K,L}$, we easily get the existence of $C_3 > 0$ only depending on $f(u)$, $M(u)$ and ψ such that $R_{K|L} \le C_3 d_{K|L} m(\sigma)$. Then using estimate (5.17) we obtain $\lim_{m\to\infty} A_{\mathcal{T}_m,\delta t_m} - B_{\mathcal{T}_m,\delta t_m} = 0$ and $\lim_{m\to\infty} A_{\mathcal{T}_m,\delta t_m} = A$, in the particular case considered here. To extend the result to the general case by density, it suffices to remark the continuity of A with respect to $M(u), f(u)$ and the uniform continuity of $A_{\mathcal{T},\delta t}$ for $L^2(\Omega\times(0,T))$ norm. The continuity of A is clear. To show uniform continuity of $A_{\mathcal{T},\delta t}$, we use first Cauchy Schwarz inequality to get the following inequalities.

$$
\begin{aligned}
\|A_{\mathcal{T},\delta t}\|^2 &\le \|\nabla\psi\|^2_{L^2(0,T,L^\infty(\Omega))}C_2\sum\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)} d_{K|L}m(K|L)(M(u)_{K|L}^{n+1})^2\\
\|A_{\mathcal{T},\delta t}\|^2 &\le \|\nabla\psi\|^2_{L^2(0,T,L^\infty(\Omega))}(M^*)^2C_2\sum\delta t\frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)} d_{K|L}m(K|L)(f(u)_{K|L}^{n+1})^2
\end{aligned}
$$

Then, since $M(u)_{K|L}^{n+1}$ belongs to the interval $[M(U_K^{n+1}), M(U_L^{n+1})]$ and $f(u)_{K|L}^{n+1} \in [f(U_K^{n+1}), f(U_L^{n+1})]$, we have

$$
\begin{aligned}
\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}(M(u)_{K|L}^{n+1})^2 d_{K|L}m(K|L) &\le 2\sum_{K\in\mathcal{T}} M(U_K^{n+1})^2(\sum_{L\in\mathcal{N}(K)} d_{K|L}m(K|L))\\
\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)}(f(u)_{K|L}^{n+1})^2 d_{K|L}m(K|L) &\le 2\sum_{K\in\mathcal{T}} f(U_K^{n+1})^2(\sum_{L\in\mathcal{N}(K)} d_{K|L}m(K|L)).
\end{aligned}
$$

By using the $\xi - regularity$ of meshes, we finally get

$$
\begin{aligned}
\|A_{\mathcal{T},\delta t}\| &\le \sqrt{C_2\xi}\|\psi\|_{L^2(0,T,L^\infty(\Omega))}\|M(u_{\mathcal{T},\delta t})\|,\\
\|A_{\mathcal{T},\delta t}\| &\le \sqrt{C_2\xi}\|\psi\|_{L^2(0,T,L^\infty(\Omega))}M^*\|f(u_{\mathcal{T},\delta t})\|,
\end{aligned}
$$

and we use the bi linearity of $A_{\mathcal{T},\delta t}$ to conclude. This completes the proof of Theorem 5.3.2. □

5.4 Numerical results

As an example of application, we made numerical experiments with the following data which are realistic in the study of oil and water flow in homogeneous porous media.

$$k_1(x) = \frac{x^3}{2}, k_2(x) = \frac{(1-x)^3}{3}$$
$$p_c(x) = -0.5\sqrt{\frac{1-x}{x}}$$

As an initial condition we take uniformly the value $u_0 = 0.5$, and take the uniform value $c = 0.8$. We represent in the following figures the behaviour of k_1, k_2, M, f and p_c. We can verify that the hypotheses of Section 5.1 are satisfied.

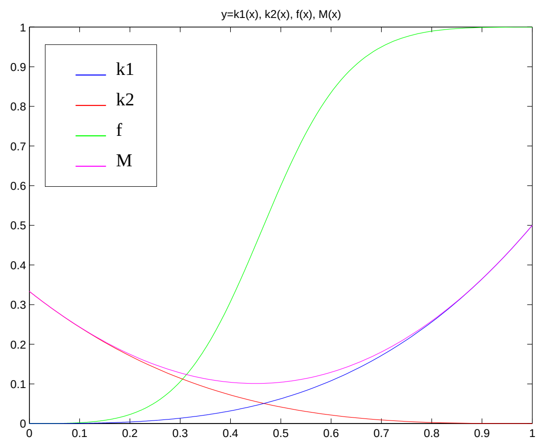


Fig 10. *Behaviour of the functions* k_1, k_2, f, M

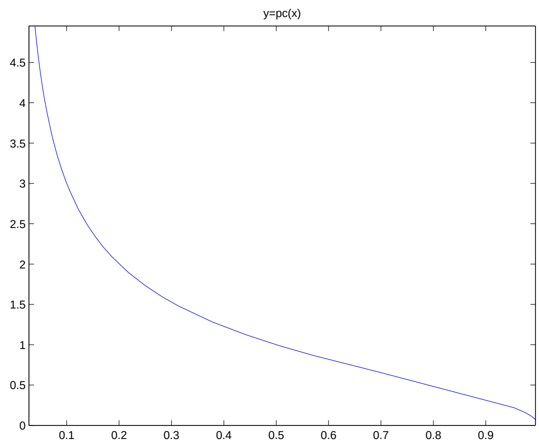


Fig 11. *Behaviour of the function* p_c.

The domain of study is the open subset $\Omega = (0,1)^2$ of $\mathbb{R}^2$. If we denote by $D_1 = \{(x,y) \in \mathbb{R}^2, (x-0.5)^2+(y-0.8)^2 \leq 0.1\}$, $D_2 = \{(x,y) \in \mathbb{R}^2, (x-0.2)^2+(y-0.2)^2 \leq 0.1\}$, $D_3 = \{(x,y) \in \mathbb{R}^2, (x-0.8)^2+(y-0.5)^2 \leq 0.1\}$, we can take sources and sinks terms as it follows :

$$\overline{s}(x,y) = 10\, \mathbb{1}_{D_1}(x,y) + 20\, \mathbb{1}_{D_2}(x,y)$$
$$\underline{s}(x,y) = 30\, \mathbb{1}_{D_3}(x,y)$$

The following figures represent the numerical results at time $t = 2.1$ and $t = 9.1$. As we could expect, the saturation in the reservoir increases in mean because $u_0 \leq c$ and the gradient of the pressure is oriented from the sources to the sink. The pressure is nearly stationary, which is not surprising since sources and sinks are stationary and the variations of the diffusion coefficient $M(u)$ is not very large in the interval $[0,1]$. The flow is more important at the beginning, when the difference between the injection saturation and the mean saturation in the reservoir is the largest.

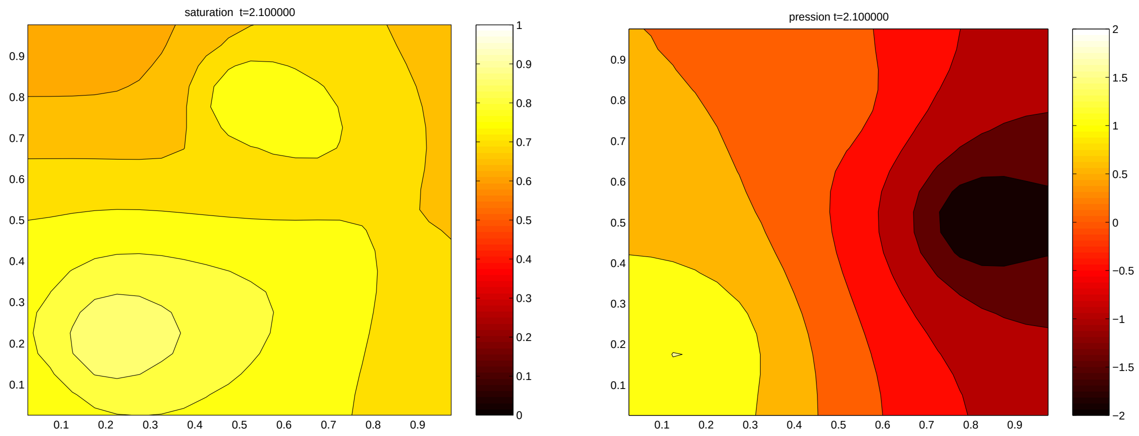


Fig 12. *Saturation and pressure at date t=2.1*

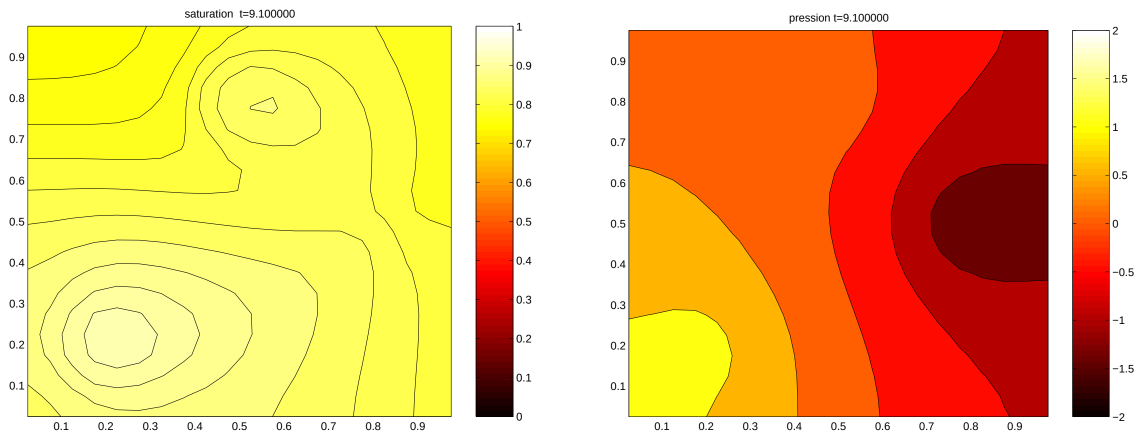


Fig 13. *Saturation and pressure at date t=9.1*

Acknowledgements. I would like to thanks M. Olhberger for helpful discussions.

Chapitre 6

Convergence d'un schéma volumes finis pour un écoulement diphasique en milieu poreux avec un décentrement phase par phase

Résumé

La modélisation des écoulements diphasiques en milieu poreux utilisés en ingénierie pétrolière conduit à un système d'équations couplées entre la saturation et la pression avec des termes de type elliptique ou parabolique dégénéré. Ces équations sont complétées par une relation entre les deux pressions, faisant intervenir la pression capillaire. Pour ce travail, nous sommes partis d'un schéma industriel basé sur une méthode volumes finis avec un décentrage phase par phase pour un modèle simplifié. Ce schéma s'avère robuste et il vérifie les contraintes physiques imposées industriellement. Pour montrer la convergence de ce schéma, nous avons prouvé dans un premier temps que le principe du maximum discret était vérifié, donc que la saturation restait dans la plage de valeurs imposées par la concentration du fluide injecté et les valeurs initiales de la saturation. Nous avons aussi montré des estimations discrètes sur la pression. Grâce à ce travail minutieux d'estimation a priori, la fin de la preuve se fait classiquement, en utilisant fortement les hypothèses.

Ce schéma a été implémenté avec succès en 1D et en 2D sur maillages rectangulaires ou triangulaires. Son comportement est très satisfaisant.

Mathematical study of a petroleum-engineering scheme

Abstract

Models of two phase flows in porous media, used in petroleum engineering, lead to a system of two coupled equations with elliptic and parabolic degenerate terms, and two unknowns, the saturation and the pressure. An industrial scheme, consisting in a finite volume method together with a phase-by-phase upstream weighting scheme, satisfies industrial constraints of robustness. This scheme is shown to satisfy some a priori estimates (the saturation is shown to remain in a fixed interval, and a discrete $L^2(0,T;H^1(\Omega))$ estimate is proved for both the pressure and a function of the saturation) which are sufficient to derive the convergence of a subsequence to a weak solution of the continuous equations as the size of the discretization tends to zero.

6.1 Introduction

Finite volumes methods have been proved to be well adapted to discretize conservative equations. For more than thirty years now, finite volume methods have been used in industry because they are cheap, simple to code and robust. The porous media problems are one of the privileged field of applications.
This success induced us to study and prove the mathematical convergence of a classical finite volume method for a simple model of two phase flow in porous media. The problem can be formulated mathematically as follows: let Ω be an open bounded subset of $\mathbb{R}^d (d \geq 0)$, $T \in \mathbb{R}^+$, find $u : \Omega \times (0,T) \to \mathbb{R}$ and $p : \Omega \times (0,T) \to \mathbb{R}$ be a solution of the coupled system:

$$
\begin{aligned}
u_t - \operatorname{div}(k_1(u)\nabla p) &= f(c)\,\overline{s} - f(u)\,\underline{s} \text{ on } \Omega \times (0,T) && (6.1.1)\\
(1-u)_t - \operatorname{div}(k_2(u)\nabla q) &= h(c)\,\overline{s} - h(u)\,\underline{s} \text{ on } \Omega \times (0,T), && (6.1.2)\\
q - p &= p_c(u) && (6.1.3)
\end{aligned}
$$

with the following Neumann boundary conditions

$$
\begin{aligned}
\nabla p \cdot \mathbf{n} = 0 \qquad & \text{on } \partial\Omega \times (0,T) && (6.1.4)\\
\nabla q \cdot \mathbf{n} = 0 \qquad & \text{on } \partial\Omega \times (0,T), && (6.1.5)
\end{aligned}
$$

the following initial condition

$$u(\cdot,0) = u_0 \text{ on } \Omega, \tag{6.1.6}$$

such that p satisfies the arbitrary homogenization equation

$$\int_\Omega p = 0 \text{ on } (0,T). \tag{6.1.7}$$

In this model, u and p are respectively the saturation and the pressure of the wetting fluid (the other fluid is called the netting fluid), k_1 and k_2 are respectively the mobilities of the wetting fluid and the mobility of the netting fluid and p_c is the capillary pressure. We suppose in particular that the physical functions k_1, k_2 and p_c only depend on the saturation u of the wetting fluid. Without any exterior action, the state of the system would be stationary because the Neumann boundary conditions (6.1.4)-(6.1.5) are homogeneous. Here, we suppose that the flow of the wetting fluid in the reservoir Ω is driven by a volumetric source with velocity $f(c)\,\overline{s}$ and a volumetric sink with rate $f(u)\,\underline{s}$ where $\overline{s}$ and $\underline{s}$ represent respectively a source term in injection and at the sinks, c is the saturation of the injected fluid, f is the reduced mobility of the wetting phase i.e.:

$$f(x) = \frac{k_1(x)}{k_1(x) + k_2(x)}. \tag{6.1.8}$$

In the same way, we denote by $h(u)$ the reduced mobility of the netting phase, i.e.:

$$h(x) = \frac{k_2(x)}{k_1(x) + k_2(x)}. \tag{6.1.9}$$

Remark 6.1.1 *As the sum of the two saturations u and $1 - u$ is equal to 1, the sum of the two reduced mobility is also equal to 1, i.e. $h(u) + f(u) = 1, \forall u \in [0, 1]$.*

In the present article we deal with the same physical problem as in Chapter 5, but the numerical method we present is different. Some extensions and justifications of this mathematical model can be found in the book of G.Chavent and J.Jaffré ([CJ86]) or in the book of J.Bear ([Bea67]). A great work have been done recently by R.Ewing and Z.Chen in a serie of articles entitled "Degenerate two phase incompressible flow"([Che97],[Che99]). These articles concern the theoritical aspects and they give usefull consequences of capillary pressure degeneracy for the regularity of the solution. Concerning the phase by phase upwinding scheme, to our knowledge, in spite of its large popularity for industry simulation codes, there is not many mathematical studies. In the article of Y. Brenier and J.Jaffré ([BJ91]), the authors give an iterative method to calculate explicitly the phase by phase upwind scheme in the case where the flow is driven by gravitationnal forces and the capillary pressure is neglected. Even if there are some similarities between this work and our study, the methods are quite different. The main reason is that our difficulties come precisely from the capillary pressure term.

The aim of our study is to show that the finite volume scheme (6.2.15)-(6.2.19) used for the approximation of the solution of (6.1.1)-(6.1.7) converges in an appropriate sense. In Section 6.2 we introduce the finite volume discretization, the numerical scheme and state the main convergence results.
Section 6.3, Section 6.4 and Section 6.5 are devoted to the proof of this result: in Section 6.3 we give a priori estimates on the approximate solution which are used to deduce the existence of an approximate solution and some compactness properties on the sequence of approximate solutions.

We make the following assumptions on the data, which we refer to in the following as hypotheses (H):

Hypotheses (H)

- (H1) Ω is a polygonal subset of $\mathbb{R}^d$, $d = 2$ or 3,
- (H2) $T > 0$ is given,
- (H3) $u_0 \in L^\infty(\Omega)$ and $0 \leq u_0(x) \leq 1$ a.e $x \in \Omega$,
- (H4) $c \in L^\infty(\Omega \times (0,T))$, $0 \leq c \leq 1$ *a.e.*,
- (H5) $\overline{s} \in L^2(\Omega \times (0,T))$, $\overline{s} \geq 0$. $\underline{s} \in L^2(\Omega \times (0,T))$, $\underline{s} \geq 0$ and $\int_\Omega \overline{s} - \underline{s} = 0$.
- (H6) $k_1 \in \mathcal{C}^1([0,1], \mathbb{R})$ and $k_1(0) = 0$,
- (H7) $k_2 \in \mathcal{C}^1([0,1], \mathbb{R})$ and $k_2(1) = 0$,

- (H8) there exists $\alpha \in \mathbb{R}$ such that $\alpha > 0$, $\frac{1}{\alpha} \geq k_1'(s) \geq \alpha$ and $\frac{1}{\alpha} \geq -k_2'(s) \geq \alpha$, for all $s \in [0,1]$,
- (H9) $p_c \in \mathcal{C}^0([0,1],\mathbb{R}) \cap \mathcal{C}^1((0,1),\mathbb{R})$ and there exists $\beta_1 < 1$ and $\beta_2 < 1$,

$$\frac{1}{\alpha s^{\beta_1}(1-s)^{\beta_2}} \geq -{p_c}'(s) \geq \frac{\alpha}{s^{\beta_1}(1-s)^{\beta_2}},$$

  for all $s \in (0,1)$,

Consequences of Hypotheses (H)

- (C1) The following properties which follow from hypotheses (H6),(H7),(H8) and (H9) are crucial to prove the convergence of the scheme (the proofs are given in Section 6.6). For all $a \in [0,1]$ and $b \in [0,1]$,

$$(b-a)\int_a^b \frac{k_2(s)}{k_1(s)+k_2(s)}\,(-{p_c}'(s))ds \leq C(\alpha,\beta_1,\beta_2)\,|\int_a^b \frac{k_1(s)k_2(s)}{k_1(s)+k_2(s)}\,(-{p_c}'(s))ds| \qquad (6.1.10)$$

  and

$$(b-a)\int_a^b \frac{k_1(s)}{k_1(s)+k_2(s)}\,(-{p_c}'(s))ds \leq C(\alpha,\beta_1,\beta_2)\,|\int_a^b \frac{k_1(s)k_2(s)}{k_1(s)+k_2(s)}\,(-{p_c}'(s))ds|$$

- (C2) For all $s \in [0,1]$, $\frac{s}{\alpha} \geq k_1(s) \geq \alpha s$, $\frac{1-s}{\alpha} \geq k_2(s) \geq \alpha(1-s)$ and $\frac{1}{\alpha} \geq k_1(s)+k_2(s) \geq \alpha$.
- (C3) Thanks to (H8), the function f defined by (6.1.8) verifies $f'(s) \geq \alpha^4$ $\forall s \in [0,1]$.
- (C4) Thanks to hypothesis (H9), the function p_c is $(1-\beta_1)$-Hölder continuous in 0 and $(1-\beta_2)$-Hölder continuous in 1.

  For the a priori estimates (see Section 6.3), we need to introduce some artificial pressures (see the "global pressure" of Chavent [CJ86]). We denote by p_g and q_g the functions defined by

$$p_g(s) = \int_0^s \frac{k_2(a)}{k_1(a)+k_2(a)}\,{p_c}'(a)da,\ \forall s \in [0,1], \qquad (6.1.11)$$

  and

$$q_g(s) = \int_0^s \frac{k_1(a)}{k_1(a)+k_2(a)}\,{p_c}'(a)da,\ \forall s \in [0,1]. \qquad (6.1.12)$$

- (C5) Thanks to hypotheses (H8) and to consequence (C4), p_g and $q_g \in \mathcal{C}^0([0,1],\mathbb{R}) \cap \mathcal{C}^1((0,1),\mathbb{R})$ The function p_g is $(1-\beta_2)$-Hölder continuous in 0 and the function q_g is $(1-\beta_1)$-Hölder continuous in 1.

- One denotes by g the function defined by

$$g(s) = -\int_0^s \frac{k_1(a)k_2(a)}{k_1(a)+k_2(a)}\,{p_c}'(a)da,\ \forall s \in [0,1]. \qquad (6.1.13)$$

  Then the function $g \in \mathcal{C}^1([0,1],\mathbb{R})$, with $g'(0)=0$, $g'(1)=0$ and $g'(s) > 0$ for all $s \in (0,1)$. One denotes by $L_g = \max_{s\in[0,1]} g'(s)$ the Lipschitz constant of g on $[0,1]$.

Definition 6.1.1 (Weak solution) *Under Hypotheses (H), we say that (u,p) is a weak solution of Problem (6.1.1)-(6.1.7) if*

$$\begin{aligned}
&p \in L^2(\Omega\times(0,T)),\\
&u \in L^\infty(\Omega\times(0,T)),\ \textit{with } 0\le u(x,t)\le 1 \textit{ for a.e. } (x,t)\in\Omega\times(0,T),\\
&p+p_g(u)\in L^2(0,T;H^1(\Omega)),\\
&g(u)\in L^2(0,T;H^1(\Omega)),
\end{aligned}$$

and for every function $\varphi\in\mathcal{D}(\Omega\times(0,T))$,

$$\int_0^T\int_\Omega u(x,t)\varphi_t(x,t) - k_1(u(x,t))\nabla p(x,t)\cdot\nabla\varphi(x,t)dxdt +$$
$$\int_0^T\int_\Omega [f(c(x,t))\,\overline{s}(x,t) - f(u(x,t))\,\underline{s}(x,t)]\varphi(x,t)dxdt + \int_\Omega u_0(x)\varphi(x,0)dx = 0$$
$$\int_0^T\int_\Omega [(1-u(x,t))\varphi_t(x,t) - k_2(u(x,t)\nabla(p+p_c(u))(x,t)\cdot\nabla\varphi(x,t)]dxdt +$$
$$\int_0^T\int_\Omega [h(c(x,t))\,\overline{s}(x,t) - h(u(x,t))\,\underline{s}(x,t)]\varphi(x,t)dxdt - \int_\Omega (1-u_0(x))\varphi(x,0)dx = 0$$
$$\int_\Omega p(x,t)dx = 0 \textit{ for a.e. } t\in(0,T)$$

Remark 6.1.2 *In the above formulation, the terms $\displaystyle\int_0^T\int_\Omega k_1(u)\nabla p\cdot\nabla\varphi$ and $\displaystyle\int_0^T\int_\Omega k_2(u)\nabla(p+p_c(u))\cdot\nabla\varphi$ are well defined thanks to the fact that :*

$$\begin{aligned}
k_1(u)\nabla p &= k_1(u)\nabla(p+p_g(u)) - \nabla g(u)\\
\textit{and } k_2(u)\nabla(p+p_c(u)) &= k_2(u)\nabla(p+p_g(u)) + \nabla g(u).
\end{aligned}$$

6.2 The finite volume scheme

6.2.1 Finite volume definitions and notations

Definition 6.2.1 (Admissible mesh of Ω) *An admissible mesh $\mathcal{T}$ of Ω is given by a set of open bounded polygonal convex subsets of Ω called control volumes and a family of points (the "centers" of control volumes) satisfying the following properties:*

1. *The closure of the union of all control volumes is $\bar\Omega$. We will denote by $|K|$ the measure of K.*
2. *For any $(K,L)\in\mathcal{T}^2$ with $K\neq L$, then $K\cap L=\emptyset$. One denotes by $\mathcal{E}\subset\mathcal{T}^2$ the set of (K,L) such that the $d-1$-Lebesgue measure of $\bar K\cap\bar L$ is positive. For $(K,L)\in\mathcal{E}$, one denotes $K|L=\bar K\cap\bar L$ and $m(K|L)$ the $d-1$-Lebesgue measure of $K|L$.*
3. *For any $K\in\mathcal{T}$, one defines $\mathcal{N}_K=\{L\in\mathcal{T},(K,L)\in\mathcal{E}\}$ and one assumes that $\partial K=\bar K\backslash K=(\bar K\cap\partial\Omega)\cup\cup_{L\in\mathcal{N}_K}K|L$.*

4. *The family of points $(x_K)_{K\in\mathcal{T}}$ is such that $x_K \in K$ (for all $K \in \mathcal{T}$) and, if $L \in \mathcal{N}_K$, it is assumed that the straight line (x_K, x_L) is orthogonal to $K|L$. We set $d_{K|L} = d(x_K, x_L)$ and $T_{K|L} = \frac{m(K|L)}{d_{K|L}}$, that is sometimes called the "transmissibility" through $K|L$.*

The problem of evolution under consideration is an evolution problem, hence we also need to discretize the time interval (0,T).

Definition 6.2.2 (Time discretization) *A time discretization of $(0, T)$ is given by an integer value N and by a strictly increasing sequence of real values $(t^n)_{n\in[\![0,N+1]\!]}$ with $t^0 = 0$ and $t^{N+1} = T$. The time steps are then defined by $\delta t^n = t^{n+1} - t^n$, for $n \in [\![0, N]\!]$.*

Then we can define a discretization of the whole domain $\Omega \times (0, T)$ by the following:

Definition 6.2.3 (Discretization of Q) *A finite volume discretization D of $\Omega \times (0, T)$ is the family $D = (\mathcal{T}, \mathcal{E}, (x_K)_{K\in\mathcal{T}}, N, (t^n)_{n\in[\![0,N]\!]})$, where $\mathcal{T}$, $\mathcal{E}$, $(x_K)_{K\in\mathcal{T}}$ is an admissible mesh of Ω in the sense of Definition 6.2.1 and N, $(t^n)_{n\in[\![0,N+1]\!]}$ is a time discretization of $(0, T)$ in the sense of Definition 6.2.2. One then sets $size(D) = \max(size(\mathcal{T}), (\delta t^n)_{n\in[\![0,N]\!]})$.*

Definition 6.2.4 (Discrete functions and notations) *Let D be a discretization of $\Omega \times (0, T)$, we denote by X_D the set of the discrete functions associated to D i.e. $X_D = \mathbb{R}^{\mathcal{T}\times[\![0,N]\!]}$. An element of X_D will be denoted with capital letters and the index D like U_D or P_D and the value at point (K, n) with the index K and the upper index n as U_K^n and P_K^{n+1}. To a discrete function U_D correspond an approximate function u_D defined almost everywhere on $\Omega \times (0, T)$ by*

$$u_D(x, t) = U_K^{n+1}, \text{ for all } (x, t) \in K \times (t^n, t^{n+1}).$$

For any function $f : \mathbb{R} \mapsto \mathbb{R}$, $f(U_D)$ will denote the discrete function $(K, n) \mapsto f(U_K^{n+1})$. If $L \in \mathcal{N}_K$, and U_D is a discrete function, we denote by $\delta_{K,L}^{n+1}(U) = U_L^{n+1} - U_K^{n+1}$. For example $\delta_{K,L}^{n+1}(f(U)) = f(U_L^{n+1}) - f(U_K^{n+1})$.

Let us now give the regularity property of discretization mesh we need to prove Theorem 6.2.1.

Definition 6.2.5 (Regularity of the mesh) *Let $\xi > 0$. A discretization D of $\Omega \times (0, T)$ is ξ-regular if*

$$\forall K \in \mathcal{T}, \sum_{L\in\mathcal{N}_K} m(K|L) d_{K|L} \leq \xi |K| \tag{6.2.14}$$

6.2.2 The coupled finite volume scheme

The finite volume scheme is obtained by writing the balance equations of the fluxes on each control volume. Let D be a discretization of $\Omega \times (0, T)$. Let us integrate equations 6.1.1-6.1.2 over each control volume K. By using the Green-Riemann formula, if Φ is a vector field, the integral of $\mathrm{div}(\Phi)$ on a control volume K is equal to the sum of the normal fluxes of Φ on the edges. Here we apply this formula to $\Phi_1 = k_1(u)\nabla p$ and $\Phi_2 = k_2(u)\nabla(p + p_c(u))$. The resulting equation is discretized with a time implicit finite diffrence scheme; the normal gradients are discretized with acenterred finite difference scheme. If we denote by $U_D = \{u_K^n\}_{n\in[\![0,N+1]\!],K\in\mathcal{T}}$ and $P_D = \{P_K^n\}_{n\in[\![1,N+1]\!],K\in\mathcal{T}}$ the discrete unknowns corresponding to u and p, the finite volume scheme that we obtain is the following set of equations:

$$U_K^0 = \frac{1}{|K|}\int_K u_0(x)dx, \text{ for all } K \in \mathcal{T}, \tag{6.2.15}$$

for all $(K,n) \in \mathcal{T} \times [\![0,N]\!]$,

$$\begin{aligned}
\frac{U_K^{n+1}-U_K^n}{\delta t^n}|K| - \sum_{L\in\mathcal{N}_K} T_{K|L}k_{1,K|L}^{n+1}\delta_{K,L}^{n+1}(P) &= |K|(f(c_K^{n+1})\,\overline{s}_K^{n+1} - f(U_K^{n+1})\,\underline{s}_K^{n+1}) && (6.2.16)\\
\frac{(1-U_K^{n+1})-(1-U_K^n)}{\delta t^n}|K| - \sum_{L\in\mathcal{N}_K} T_{K|L}k_{2,K|L}^{n+1}\delta_{K,L}^{n+1}(Q) &= |K|(h(c_K^{n+1})\,\overline{s}_K^{n+1} - h(U_K^{n+1})\,\underline{s}_K^{n+1}) && (6.2.17)\\
Q_K^{n+1} - P_K^{n+1} &= p_c(U_K^{n+1}) && (6.2.18)
\end{aligned}$$

and

$$\sum_{K\in\mathcal{T}} |K|P_K^{n+1} = 0, \text{ for all } n \in [\![0,N]\!], \tag{6.2.19}$$

where

- c_K^{n+1} is the mean of c over the the time-space cell $K \times (t^n, t^{n+1})$,
- $\overline{s}_K^{n+1}$ and $\underline{s}_K^{n+1}$ denote the mean of $\overline{s}$ and $\underline{s}$ over the time-space cell $K \times (t_n, t_{n+1})$,
- The upwind discretization of $k_1(u)$ (or $k_2(u)$) for the interface $K|L$ which is respectively denoted by $k_{1,K|L}^{n+1}$ (or $k_{2,K|L}^{n+1}$) is defined in the following.

Let $\mathcal{E}_1^{n+1}$ and $\mathcal{E}_2^{n+1}$ be two subsets of $\mathcal{E}$ such that:

1. $\{(K,L) \in \mathcal{E},\ P_L^{n+1} - P_K^{n+1} < 0\} \subset \mathcal{E}_1^{n+1} \subset \{(K,L) \in \mathcal{E},\ P_L^{n+1} - P_K^{n+1} \le 0\}$
2. $\{(K,L) \in \mathcal{E},\ Q_L^{n+1} - Q_K^{n+1} < 0\} \subset \mathcal{E}_2^{n+1} \subset \{(K,L) \in \mathcal{E},\ Q_L^{n+1} - Q_K^{n+1} \le 0\}$
3. if $(K,L) \in \mathcal{E}_1^{n+1}$ then $(L,K) \notin \mathcal{E}_1^{n+1}$.
4. if $(K,L) \in \mathcal{E}_2^{n+1}$ then $(K,L) \notin \mathcal{E}_2^{n+1}$.

We define

$$U_{1,K|L}^{n+1} = \begin{cases} U_K^{n+1} \text{ if } (K,L) \in \mathcal{E}_1^{n+1} \\ U_L^{n+1} \text{ otherwise} \end{cases}, \qquad U_{2,K|L}^{n+1} = \begin{cases} U_K^{n+1} \text{ if } (K,L) \in \mathcal{E}_2^{n+1} \\ U_L^{n+1} \text{ otherwise} \end{cases}. \tag{6.2.20}$$

and

$$k_{1,K|L}^{n+1} = k_1(U_{1,K|L}^{n+1}), \tag{6.2.21}$$
$$k_{2,K|L}^{n+1} = k_2(U_{2,K|L}^{n+1}). \tag{6.2.22}$$

Remark 6.2.1 *The formulas (6.2.20) express a phase by phase upstream choice : the value of the reduced mobilities of each phase on the edge (K,L) is determined by the sign of the discrete of the pressure. The null flux edges (the edges (K,L) where the values of the pressure in K and L is the same) have to be specified as elements of $\mathcal{E}_1^{n+1}$ or its complementary because of technical reasons in the proof of Proposition 6.3.38.*

We show below (see proposition 6.3.4)) that there exists at least a solution to this scheme. From this discrete solution, we build an approximate solution (u_D, p_D) defined almost everywhere on $\Omega \times (0, T)$ by (Cf Definition 6.2.4 :

$$\begin{aligned} u_D(x,t) &= U_K^{n+1}, \qquad \forall x \in K, \forall t \in (t_n, t_{n+1}), \\ p_D(x,t) &= P_K^{n+1}, \qquad \forall x \in K, \forall t \in (t_n, t_{n+1}). \end{aligned}$$

Remark 6.2.2 *Practically, we solve a non-linear fixed point problem with a multidimensional Newton method. Numerical experiments show that if the time step is adequately chosen, the Newton procedure converges with a small number of iterations hence this scheme is cheaper than the analogous explicit one.*

We may now state the main convergence result.

Theorem 6.2.1 *Assume Hypothesis (H) are satisfied. Let $\{D_n\}_{n\in\mathbb{N}}$ be a sequence of discretizations of $\Omega \times (0, T)$ in the sense of Definition 6.2.3, uniformly satisfying the ξ-regularity property (6.2.14) and such that $\lim_{n\to\infty} size(D_n) = 0$. Let (u_{D_n}, p_{D_n}) be the approximate solutions corresponding to D_n. Then (u_{D_n}, p_{D_n}) converges to a weak solution (u, p) of (6.1.1)-(6.1.7) .*

Remark 6.2.3 *If we try to write this scheme as a coupling of a discrete parabolic equation on U_D and a discrete elliptic equation on P_D, using the global pressure formulation of Chavent [CJ86], we do not obtain a monotone flux scheme. Thus the method used to obtain a priori estimates in [Mic01] cannot be used there.*

6.3 Discrete a priori estimates

In this section, we develop the first part of the proof of Theorem 6.2.1. The method used to prove these estimates is quite different from the one used in the previous related papers ([Mic01], [EGHNS98]), and is interesting for its own sake. Since the continuous case is the guideline of our reasoning, we give a sketch of the ideas that we follow before each proof.

6.3.1 The maximum principle

We show that the phase by phase upstream choice yields the stability of the scheme.

Proposition 6.3.1 (The maximum principle) *Assume that hypotheses (H) are fulfilled and (U_D, P_D) is a solution of the finite volume scheme (6.2.15)-(6.2.19) then we have the following maximum principle :*

$$0 \le U_K^n \le 1, \qquad \forall K \in \mathcal{T}, \forall n \in [\![0, N+1]\!]. \tag{6.3.23}$$

Proof. By symmetry, we only need to prove the right part of Inequality (6.3.23). By contradiction, let us assume that the maximum value of U_D on $\mathcal{T} \times [\![0, N+1]\!]$ is larger than 1. Then it cannot be attained for U_K^0, since the initial condition (6.2.15) clearly implies that $U_K^0 \le 1$, so there exists $n \ge 0$ such that the maximum value of U_D is U_K^{n+1}. If n is chosen minimal, $U_K^{n+1} > U_K^n$ so by using (6.2.16) and (6.2.17) we have

$$\sum_{L\in\mathcal{N}_K} T_{K|L}k_{1,K|L}^{n+1}\delta_{K,L}^{n+1}(P)+|K|(f(c_K^{n+1})\,\overline{s}_K^{n+1}-f(U_K^{n+1})\,\underline{s}_K^{n+1}) \quad > \quad 0, \tag{6.3.24}$$

$$-\sum_{L\in\mathcal{N}_K} T_{K|L}k_{2,K|L}^{n+1}\delta_{K,L}^{n+1}(Q)-|K|(h(c_K^{n+1})\,\overline{s}_K^{n+1}-h(U_K^{n+1})\,\underline{s}_K^{n+1}) \quad > \quad 0. \tag{6.3.25}$$

By Definition (6.2.21)-(6.2.22) of the upwind approximation, the terms $T_{K|L}k_{1,K|L}^{n+1}\delta_{K,L}^{n+1}(P)$ and $-T_{K|L}k_{2,K|L}^{n+1}\delta_{K,L}^{n+1}(Q)$ are nondecreasing with respect to U_L^{n+1} so that, Inequalities (6.3.24) and (6.3.25) remains valid replacing U_L^{n+1} by U_K^{n+1}. Thus we obtain :

$$k_1(U_K^{n+1})\sum_{L\in\mathcal{N}_K} T_{K|L}\delta_{K,L}^{n+1}(P)+|K|(f(c_K^{n+1})\,\overline{s}_K^{n+1}-f(U_K^{n+1})\,\underline{s}_K^{n+1})>0,$$
$$-k_2(U_K^{n+1})\sum_{L\in\mathcal{N}_K} T_{K|L}\delta_{K,L}^{n+1}(Q)-|K|(h(c_K^{n+1})\,\overline{s}_K^{n+1}-h(U_K^{n+1})\,\underline{s}_K^{n+1})>0.$$

And by hypothesis, since p_c is non increasing, $\delta_{K,L}^{n+1}(Q)\geq\delta_{K,L}^{n+1}(P)$, so we also have

$$-k_2(U_K^{n+1})\sum_{L\in\mathcal{N}_K} T_{K|L}\delta_{K,L}^{n+1}(P)-|K|(h(c_K^{n+1})\,\overline{s}_K^{n+1}-h(U_K^{n+1})\,\underline{s}_K^{n+1})>0. \tag{6.3.26}$$

Now let us multiply (6.3.26) by $k_2(U_K^{n+1})$, (6.3.26) by $k_1(U_K^{n+1})$ and sum the two resulting inequalities. This yields:

$$(k_2(U_K^{n+1})f(c_K^{n+1})-k_1(U_K^{n+1})h(c_K^{n+1}))|K|\,\overline{s}_K^{n+1}>0 \tag{6.3.27}$$

Now, since k_1 is non increasing and k_2 is nondecreasing, the left hand side in Inequality (6.3.27) is non-increasing with respect to U_K^{n+1} and is equal to zero if $U_K^{n+1}=c_K^{n+1}$. So it is in contradiction with the hypothesis $U_K^{n+1}>1$ since $c_K^{n+1}\leq 1$.
□

Remark 6.3.1 *The maximum principle of Proposition 6.3.23 is crucial in applications since the functions $k_1(u)$ and $k_2(u)$ are only defined for $u\in[0,1]$.*

6.3.2 Estimates on the pressure

The following proposition is a preliminary to the proof of the estimates given in Proposition 6.3.3.

Proposition 6.3.2 (Preliminary) *Under Hypotheses (H), one denotes by p_g the function defined, for all $s\in[0,1]$, by $p_g(s)=\displaystyle\int_0^s \frac{k_2(a)}{k_1(a)+k_2(a)}\,{p_c}'(a)da$. Let D be a finite volume discretization of $\Omega\times(0,T)$ in the sense of Definition 6.2.3 and let (U_D,P_D) be a solution of (6.2.15)-(6.2.19) . Then the following inequalities hold:*

$$k_{1,K|L}^{n+1}+k_{2,K|L}^{n+1}\geq\alpha,\ \forall(K,L)\in\mathcal{E},\ \forall n\in[\![0,N]\!], \tag{6.3.28}$$

and $\forall(K,L)\in\mathcal{E},\ \forall n\in[\![0,N]\!],$

$$\alpha\left(\delta_{K,L}^{n+1}(P)+\delta_{K,L}^{n+1}(p_g(U))\right)^2 \le k_{1,K|L}^{n+1}\left(\delta_{K,L}^{n+1}(P)\right)^2 + k_{2,K|L}^{n+1}\left(\delta_{K,L}^{n+1}(Q)\right)^2 \tag{6.3.29}$$

Proof. To make things clear, we first sketch the proof of the continuous equivalent of (6.3.28) and (6.3.29). We first give in Step 2 the proof in the discrete setting, which is adapted from the continuous one.

Step 1. Proof of the continuous equivalent of (6.3.28) and (6.3.29).
The fact that $k_1(u)+k_2(u)\ge\alpha$ is a direct consequence of hypotheses (H8). This signifies that the total mobility cannot be reduced to zero in the fluid.

Now the continuous equivalent of Inequality (6.3.29) writes

$$\alpha\nabla(p+p_g(u)))^2 \le k_1(u)(\nabla p)^2 + k_2(u)(\nabla q)^2 \tag{6.3.30}$$

By definition, $f(u)+h(u)=1$, so by convexity,

$$\begin{aligned}(\nabla(p+p_g(u)))^2 &= (f(u)\nabla p + h(u)\nabla q)^2\\ &\le f(u)(\nabla p)^2 + h(u)(\nabla q)^2\\ &\le \frac{1}{k_1(u)+k_2(u)}(k_1(u)(\nabla p)^2 + k_2(u)(\nabla q)^2),\end{aligned}$$

and hence (6.3.30) follows from the fact that $k_1(u)+k_2(u)\ge\alpha$.

Step 2. Proof in the discrete setting
In order to prove (6.3.28), we shall study the different possible cases separetly.

If $(K,L)\in\mathcal{E}_1^{n+1}\cap\mathcal{E}_2^{n+1}$ or $(K,L)\notin\mathcal{E}_1^{n+1}\cup\mathcal{E}_2^{n+1}$, $U_{1,K|L}^{n+1}=U_{2,K|L}^{n+1}$ so (6.3.28) is an immediate consequence of Hypotheses (H).

If $(K,L)\in\mathcal{E}_1^{n+1}$ and $(K,L)\notin\mathcal{E}_2^{n+1}$, then $U_{1,K|L}^{n+1}=U_K^{n+1}$, $U_{2,K|L}^{n+1}=U_L^{n+1}$ and $\delta_{K,L}^{n+1}(p_c(U))=\delta_{K,L}^{n+1}(Q)-\delta_{K,L}^{n+1}(P)\ge 0$, which yields $U_K^{n+1}\ge U_L^{n+1}$. Therefore

$$\begin{aligned}k_{1,K|L}^{n+1}+k_{2,K|L}^{n+1} \ge k_1(U_K^{n+1})+k_2(U_K^{n+1})\ge\alpha\\ k_{1,K|L}^{n+1}+k_{2,K|L}^{n+1} \ge k_1(U_L^{n+1})+k_2(U_L^{n+1})\ge\alpha\end{aligned}$$

The case $(K,L)\notin\mathcal{E}_1^{n+1}$ and $(K,L)\in\mathcal{E}_2^{n+1}$ is similar.

Remark 6.3.2 *We showed that if the upwind choice is different for the two equations, then*

$$k_{1,K|L}^{n+1} = \max_{[U_K^{n+1},U_L^{n+1}]} k_1 \quad \textit{and} \quad k_{2,K|L}^{n+1} = \max_{[U_K^{n+1},U_L^{n+1}]} k_2.$$

The proof of (6.3.29) will be obtained in two steps.

If $(K,L) \in \mathcal{E}_1^{n+1}$ and $(K,L) \notin \mathcal{E}_2^{n+1}$. By definition of p_g there exists some $a_0 \in [U_L^{n+1}, U_K^{n+1}]$ such that $\delta_{K,L}^{n+1}(p_g(U)) = h(a_0)\delta_{K,L}^{n+1}(p_c(U))$, so by using $f+h=1$ and $f \le 1$, $h \le 1$ we get

$$\begin{aligned}(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_g(U)))^2 \quad &= (f(a_0)\delta_{K,L}^{n+1}(P) + h(a_0)(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_c(U))))^2 \\ &\le f(a_0)(\delta_{K,L}^{n+1}(P))^2 + h(a_0)(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_c(U)))^2,\end{aligned}$$

Now, using Remark 6.3.2, we have $k_1(a_0) \le k_{1,K|L}^{n+1}$ and $k_2(a_0) \le k_{2,K|L}^{n+1}$ so that we may write

$$(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_g(U)))^2 \quad \le \tfrac{k_{1,K|L}^{n+1}}{k_1(a_0)+k_2(a_0)}(\delta_{K,L}^{n+1}(P))^2 + \tfrac{k_{2,K|L}^{n+1}}{k_1(a_0)+k_2(a_0)}(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_c(U)))^2,$$

which gives a fortiori (6.3.29). The case $(K,L) \notin \mathcal{E}_1^{n+1}$ and $(K,L) \in \mathcal{E}_2^{n+1}$ is similar.

Let us now deal with the other case. If $(K,L) \in \mathcal{E}_1^{n+1}$ and $(K,L) \in \mathcal{E}_2^{n+1}$ then $k_{1,K|L}^{n+1} = k_1(U_K^{n+1})$ and $k_{2,K|L}^{n+1} = k_2(U_K^{n+1})$. We then remark that, since the function h is nondecreasing and p_c is non increasing, the following inequality holds:

$$\begin{aligned}&k_2(U_K^{n+1})\delta_{K,L}^{n+1}(p_c(U)) - (k_1(U_K^{n+1}) + k_2(U_K^{n+1}))\delta_{K,L}^{n+1}(p_g(U)) = \\ &\quad (k_1(U_K^{n+1}) + k_2(U_K^{n+1}))\int_{U_K^{n+1}}^{U_L^{n+1}} (h(U_K^{n+1}) - h(a))p_c'(a)da \le 0\end{aligned}$$

one then gets

$$\begin{aligned}[\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_c(U))][k_2(U_K^{n+1})\delta_{K,L}^{n+1}(p_c(U)) - (k_1(U_K^{n+1}) + k_2(U_K^{n+1}))\delta_{K,L}^{n+1}(p_g(U))] \ge 0 \\ \delta_{K,L}^{n+1}(P)[k_2(U_K^{n+1})\delta_{K,L}^{n+1}(p_c(U)) - (k_1(U_K^{n+1}) + k_2(U_K^{n+1}))\delta_{K,L}^{n+1}(p_g(U))] \ge 0\end{aligned}$$

This leads by adding the two inequalities to

$$\begin{aligned}&2k_2(U_K^{n+1})\delta_{K,L}^{n+1}(P)\delta_{K,L}^{n+1}(p_c(U)) + k_2(U_K^{n+1})(\delta_{K,L}^{n+1}(p_c(U)))^2 \ge \\ &(k_1(U_K^{n+1}) + k_2(U_K^{n+1}))[2\delta_{K,L}^{n+1}(P)\delta_{K,L}^{n+1}(p_g(U)) + \delta_{K,L}^{n+1}(p_c(U))\delta_{K,L}^{n+1}(p_g(U))] \ge \\ &(k_1(U_K^{n+1}) + k_2(U_K^{n+1}))[2\delta_{K,L}^{n+1}(P)\delta_{K,L}^{n+1}(p_g(U)) + (\delta_{K,L}^{n+1}(p_g(U)))^2].\end{aligned}$$

The previous inequality gives

$$\begin{aligned}k_1(U_K^{n+1})(\delta_{K,L}^{n+1}(P))^2 \quad &+ \quad k_2(U_K^{n+1})(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_c(U)))^2 \\ &\ge \quad (k_1(U_K^{n+1}) + k_2(U_K^{n+1}))(\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_g(U)))^2,\end{aligned}$$

which is (6.3.29) in that case. The case $(K,L) \notin \mathcal{E}_1^{n+1}$ and $(K,L) \notin \mathcal{E}_2^{n+1}$ is similar. □

We can now state the following property.

Proposition 6.3.3 (Pressure estimates) *Under hypotheses (H), one denotes by p_g the function defined, for all $s \in [0,1]$, by $p_g(s) = \displaystyle\int_0^s \frac{k_2(a)}{k_1(a)+k_2(a)}\, {p_c}'(a)da$. Let D be a finite volume discretization of $\Omega \times (0,T)$ using the notations of Definition 6.2.3 and let (U_D, P_D) be a solution of the finite volume scheme (6.2.15)-(6.2.19) , there exists $C_1 > 0$, which only depends on k_1, k_2, p_c, Ω, T, u_0, $\overline{s}$, $\underline{s}$, and not on D, such that the following discrete $L^2(0,T;H^1(\Omega))$ estimates hold*

$$\frac{1}{2}\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} (\delta_{K,L}^{n+1}(P))^2 \le C_1\,, \tag{6.3.31}$$

$$\frac{1}{2}\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} T_{K|L} k_{2,K|L}^{n+1} (\delta_{K,L}^{n+1}(Q))^2 \le C_1 \tag{6.3.32}$$

and

$$\frac{1}{2}\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} T_{K|L} (\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_g(U)))^2 \le C_1\,. \tag{6.3.33}$$

Proof. Before proving this estimate, we shall give in Step 1 a formal proof in the continuous case to underline the main ideas.

Step 1. Proof in the continuous case

Suppose that u and p are regular functions that satisfy the coupled system of equations and let us multiply the first equation by p and the second one by q. Then making the difference between the two equations and integrate over $\Omega \times (0,T)$, we get :

$$\int_0^T\!\!\int_\Omega u_t(-p_c(u)) + \int_0^T\!\!\int_\Omega k_1(u)(\nabla p)^2 + k_2(u)(\nabla q)^2 = \int_0^T\!\!\int_\Omega (f(c)\,\overline{s} - f(u)\,\underline{s})p - (h(c)\,\overline{s} - h(u)\,\underline{s})q) \tag{6.3.34}$$

Let g_c be a primitive of $-p_c$. Then

$$\int_0^T\!\!\int_\Omega u_t(-p_c(u)) = \int_\Omega g_c(u(T)) - g_c(u_0)$$

which is bounded by using the maximum principle.

The second term is positive and greater than $\alpha \displaystyle\int_0^T\!\!\int_\Omega (\nabla(p+p_g(u)))^2$ by using the preliminary proposition. Then this equation gives a bound for $\nabla(p+p_g(u))$ in $L^2(\Omega\times(0,T))$ provided that we control the second term. Since $\nabla p + p_g(u) = \nabla q + q_g(u)$, we may use $p+p_g(u)$ or $q+q_g(u)$ when it is necessary. Besides, the following equations holds:

$$\begin{aligned}
(f(c)\,\overline{s} - f(u)\,\underline{s})p - (h(c)\,\overline{s} - h(u)\,\underline{s})q) \;&=\; (f(c)\,\overline{s} - f(u)\,\underline{s})(p+p_g(u)) \\
&\quad -(h(c)\,\overline{s} - h(u)\,\underline{s})(q+q_g(u)) \\
&\quad +(f(c)\,\overline{s} - f(u)\,\underline{s})p_g(u) + (h(c)\,\overline{s} - h(u)\,\underline{s})q_g(u) \\
&=\; (\,\overline{s} - \,\underline{s})(p+p_g(u)) + (h(c)\,\overline{s} - h(u)\,\underline{s})C_g \\
&\quad +p_g(u)(f(c)\,\overline{s} - f(u)\,\underline{s}) + q_g(u)(h(c)\,\overline{s} - h(u)\,\underline{s})
\end{aligned}$$

so by the Poincaré inequality and using the fact that p_g and q_g are continuous functions, we obtain

$$|\int_0^T\int_\Omega (f(c)\,\overline{s} - f(u)\,\underline{s})p - (h(c)\,\overline{s} - h(u)\,\underline{s})q| \le C_1\|p + p_g(u)\|_{L^2(Q)} + C_2$$

Then we get a bound on $\nabla(p+p_g)^2$ in $L^1(\Omega \times (0,T))$ *i.e.* a $L^2(0,T,H^1(\Omega))$ bound on $p+p_g$. Analoguous bounds on $k_1(u)\nabla p^2$ and $k_2(u)\nabla q^2$ then easily come from equation (6.3.34). This completes the proof in the continuous case.

Step 2. Proof of (6.3.31)-(6.3.33) (discrete case)
In the following proof, we denote by C_i various real values which only depend on k_1, k_2, p_c, Ω, T, u_0, $\overline{s}$, $\underline{s}$, and not on D. Let us multiply (6.2.16) by $\delta t^n P_K^{n+1}$ and (6.2.17) by $\delta t^n Q_K^{n+1}$ and sum the two equations thus obtained. Next we sum the result over $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$. We obtain

$$\begin{aligned}
- & \sum_{n=0}^{N}\sum_{K\in\mathcal{T}} |K|(U_K^{n+1} - U_K^n)p_c(U_K^{n+1}) \\
+ & \frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K} T_{K|L}k_{1,K|L}^{n+1}|\delta_{K,L}^{n+1}(P)|^2 \\
+ & \frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K} T_{K|L}k_{2,K|L}^{n+1}|\delta_{K,L}^{n+1}(Q)|^2 \\
\le & \; C_2 + \sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} |K|(\,\overline{s}_K^{n+1} - \underline{s}_K^{n+1})P_K^{n+1}.
\end{aligned} \tag{6.3.35}$$

We introduce the function $g_c \in \mathcal{C}^1([0,1], \mathbb{R}^+)$, defined by

$$g_c(s) = \int_s^1 p_c(a)da, \;\; \forall s \in [0,1].$$

Since p_c is a decreasing function, the function g_c is convex. We thus get

$$-(U_K^{n+1} - U_K^n)p_c(U_K^{n+1}) \ge g_c(U_K^{n+1}) - g_c(U_K^n), \;\; \forall K \in \mathcal{T}, \;\forall n \in \mathbb{N}. \tag{6.3.36}$$

Besides, thanks to the discrete Poincaré inequality (see [CVV99]), which is available here since

$$\left|\sum_{K\in\mathcal{T}} |K|(P_K^{n+1} + p_g(U_K^{n+1}))\right| = \left|\sum_{K\in\mathcal{T}} |K|p_g(U_K^{n+1})\right| \le C_3\,,$$

and the $L^\infty(\Omega \times (0,T))$ estimate (6.3.23), we get the existence of C_4 and of C_5 such that

$$\begin{aligned}
\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} |K|(\,\overline{s}_K^{n+1} - \underline{s}_K^{n+1})P_K^{n+1} \;\; = \;\; & \sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} |K|(\,\overline{s}_K^{n+1} - \underline{s}_K^{n+1})(P_K^{n+1} + p_g(U_K^{n+1})) \\
& -\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} |K|(\,\overline{s}_K^{n+1} - \underline{s}_K^{n+1})p_g(U_K^{n+1})
\end{aligned}$$

$$\leq \left(\frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K} T_{K|L}(\delta^{n+1}_{K,L}(P)+\delta^{n+1}_{K,L}(p_g(U)))^2\right)^{1/2} + C_5\,.$$

It leads by Young's inequality to the existence of C_6 such that

$$\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}|K|(\overline{s}^{n+1}_K - \underline{s}^{n+1}_K)P^{n+1}_K \leq \frac{\alpha}{4}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K} T_{K|L}(\delta^{n+1}_{K,L}(P)+\delta^{n+1}_{K,L}(p_g(U)))^2 + C_6\,. \quad (6.3.37)$$

Inequalities (6.3.29),(6.3.35), (6.3.36) and (6.3.37) give (6.3.33). □

6.3.3 Existence of a discrete solution

We prove here the existence of a solution to the scheme, which is a consequence of Leray-Schauder fixed point theorem. The idea of the proof is the following: if we can modify continuously the scheme to obtain a linear system which has a solution and if the modification preserves in the same time the estimates, then the scheme also has a solution.

Proposition 6.3.4 *Under Hypothesis (H), there exists a solution (U_D, P_D) to the scheme (6.2.15)-(6.2.19)*

Proof.
We define the vector space of discrete solutions E_D by

$$E_D = \mathbb{R}^{\mathcal{T}\times[\![0,N+1]\!]} \times \mathbb{R}^{\mathcal{T}\times[\![1,N+1]\!]}.$$

We define a continuous application $\mathcal{F} : [0,1]\times E_D \to E_D$ by $\mathcal{F}(t,(U_D,P_D)) = (A_D,B_D)$, where

$$A^0_K = U^0_K - \frac{1}{|K|}\int_K u_0(x)dx, \text{ for all } K\in\mathcal{T}.$$

forall $(K,n)\in\mathcal{T}\times[\![0,N]\!]$,

$$A^{n+1}_K = \frac{U^{n+1}_K - U^n_K}{\delta t^n}|K| - \sum_{L\in\mathcal{N}_K} T_{K|L}{k^t_1}^{n+1}_{K|L}\delta^{n+1}_{K,L}(P) - |K|(f^t(c^{n+1}_K)t\,\overline{s}^{n+1}_K + f^t(U^{n+1}_K)t\,\underline{s}^{n+1}_K)$$

for all $K\in\mathcal{T}, K\neq K_0$, for all $n\in[\![0,N]\!]$,

$$B^{n+1}_K = \frac{U^n_K - U^{n+1}_K}{\delta t^n}|K| - \sum_{L\in\mathcal{N}_K} T_{K|L}{k^t_2}^{n+1}_{K|L}(\delta^{n+1}_{K,L}(P)+\delta^{n+1}_{K,L}({p_c}^t(U))) - |K|(h^t(c^{n+1}_K)t\,\overline{s}^{n+1}_K + h^t(U^{n+1}_K))t\,\underline{s}^{n+1}_K)$$

and for the control volume K_0,

$$B^{n+1}_{K_0} = \sum_{K\in\mathcal{T}}|K|P^{n+1}_K, \text{ for all } n\in[\![0,N]\!],$$

where K_0 is an arbitrary control volume of $\mathcal{T}$ used to fix the arbitrary condition on the pressure due to the Neumann boundary conditions. Here u_0^t, k_1^t, k_2^t, f^t and ${p_c}^t$ are continuous modifications of u_0, k_1, k_2, f and p_c that preserve the properties used to obtain the maximum principle and the pressure estimates. Precisely, we take $x_0 \in [0,1]$, we denote by H^t the function defined by $H^t(x) = tx + (1-t)(x_0)$ and we choose $k_1^t = k_1 \circ H^t$, $k_2^t = k_2 \circ H^t$, $f^t = f \circ H^t$ and ${p_c}^t = p_c \circ H^t$. Then, the definition of $k_1^t(u)_{K|L}^{n+1}$ and $k_2^t(u)_{K|L}^{n+1}$ is the analogue of definition of $k_{1,K|L}^{n+1}$ and $k_{2,K|L}^{n+1}$ by (6.2.21),(6.2.22) with k_1^t, k_2^t and Q^t instead of k_1, k_2 and Q, with $\delta_{K,L}^{n+1}(Q^t) = \delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}({p_c}^t(U))$.

Let us now complete the proof. First of all, $\mathcal{F}(0,\cdot)$ is clearly an affine function. Moreover $\mathcal{F}(t,(U_D,P_D)) = 0$ if and only if (U_D,P_D) is a solution to the scheme with functions u_0^t, k_1^t, k_2^t, f^t, ${p_c}^t$, $t\,\overline{s}$, and $t\,\underline{s}$ so by using a priori estimates and Poincaré discrete inequality, we get a bound on U_D and P_D independent of t. The function $\mathcal{F}$ is continuous. Indeed, the ambiguous terms, which are exactly the terms corresponding to the phase by phase upstreaming can be expressed differently with the help of the continuous functions $x \mapsto x^+$ and $x \mapsto x^-$ as in the following :

$$\begin{aligned} {k^t_1}_{K|L}^{n+1}\delta_{K,L}^{n+1}(P) &= k_1^t(U_K^{n+1})(\delta_{K,L}^{n+1}(P))^+ - k_1^t(U_L^{n+1})(\delta_{K,L}^{n+1}(P))^- \\ {k^t_2}_{K|L}^{n+1}\delta_{K,L}^{n+1}(Q^t) &= k_2^t(U_K^{n+1})(\delta_{K,L}^{n+1}(Q^t))^+ - k_2^t(U_L^{n+1})(\delta_{K,L}^{n+1}(Q^t))^-. \end{aligned}$$

By Leray-Shauder's topological degree theorem (see [Dei85]). If X is a ball with a sufficiently large radius in E_D, the equation $\mathcal{F}(t,(U_D,P_D)) = 0$ has no solution on the boundary of X, so that

$$degree(\mathcal{F}(1,\cdot),X) = degree(\mathcal{F}(0,\cdot),X) = det(\mathcal{F}(0,\cdot)) \neq 0,$$

and thus there exists at least a solution to the scheme.

Remark 6.3.3 *The Leray-Shauder topological degree theorem was used to prove existence of a solution to a finite volume scheme for the first time in [EGGH98]. The proof is quitly different here because the source terms depends on the functions k_1 and k_2 which are present in the left hand side of the coupled system that we study.*

Remark 6.3.4 *The linear part of the application $\mathcal{F}(0,\cdot)$ is clearly an isomorphism since if $\ker(\mathcal{F}(0,\cdot)) \neq \{0\}$, the subset $\{x \in E_D, \mathcal{F}(0,x) = 0\}$ would intersect ∂X.*

Remark 6.3.5 *To show that $\mathcal{F}(t,(U_D,P_D)) = 0$ if and only if (U_D,P_D) is a solution to the scheme, we must recall that $\sum_{K\in\mathcal{T}} \overline{s}_K^{n+1} - \underline{s}_K^{n+1} = 0$ because by hypothesis, $\int_\Omega \overline{s} - \underline{s} = 0$. Hence the equation for $K = K_0$ can be obtained by summing all the equations corresponding to the other control volumes.*

□

6.3.4 Estimates on $g(u)$

The following estimate is used below to prove a compactness property on U_D. The analogue of the proof in the continuous case is not very difficult, but it uses strongly the symmetry of the system. The discrete proof is somewhat more complicated because of the phase to phase upstream weighting.

Proposition 6.3.5 *Under Hypotheses (H), one denotes by g the function defined, for all $s \in [0,1]$, by $g(s) = -\int_0^s \frac{k_1(a)k_2(a)}{k_1(a)+k_2(a)}\, {p_c}'(a)da$. Let D be a finite volume discretization of $\Omega \times (0,T)$ in the sense of Definition 6.2.3 and let (U_D, P_D) be a solution of the finite volume scheme (6.2.15)-(6.2.19) .*
Then there exists C_7, which only depends on k_1, k_2, p_c, Ω, T, u_0, $\overline{s}$, $\underline{s}$, and not on D, such that the following discrete $L^2(0,T;H^1(\Omega))$ estimate holds :

$$\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} T_{K|L}\delta^{n+1}_{K,L}(g(U))\delta^{n+1}_{K,L}(f(U)) \le C_7 \tag{6.3.38}$$

Proof.
Step 1. Proof in the continuous case .
As we did for pressure estimates, we give here the ideas of the proof in the continuous case, assuming that (u,p) is a regular solution. The continuous estimate to (6.3.38) writes

$$\int_0^T \int_\Omega \nabla g(u) \nabla f(u) \le C \tag{6.3.39}$$

To preserve the symmetry of the system, we multiply the first equation by $f(u)$ and the second equation by $h(u)$. Summing the two equations we obtain:

$$\begin{aligned}\int_0^T\int_\Omega u_t(f(u)-h(u)) + \operatorname{div}(k_1(u)\nabla p)f(u) + \operatorname{div}(k_2(u)\nabla q)h(u) \quad &= \int_0^T\int_\Omega f(u)(f(c)\,\overline{s} - f(u)\,\underline{s}) \\ &+ \int_0^T\int_\Omega h(u)(h(c)\,\overline{s} - h(u)\,\underline{s}).\end{aligned}$$

Let us introduce the total velocity flow F which writes

$$F \;=\; k_1(u)\nabla p + k_2(u)\nabla q$$

Remarking that $k_2(u){p_c}'(u)\nabla u = (k_1(u)+k_2(u))\,\nabla p_g(u)$, one has

$$F = (k_1(u)+k_2(u))\nabla(p+p_g(u)) \text{ and } F = (k_1(u)+k_2(u))\nabla(q-q_g(u)),$$

so that

$$k_1(u)\nabla p = f(u)F - (k_1(u)\nabla p_g(u)) \text{ and } k_2(u)\nabla q = h(u)F - (k_2(u)\nabla q_g(u)),$$

and by definition of p_g, q_g, and g (6.1.11)-(6.1.13) , one also has

$$k_2(u)\nabla p_g(u) = k_2(u)\nabla q_g(u) = -\nabla g(u).$$

This gives:

$$\begin{aligned}\int_0^T\int_\Omega u_t(f(u)-h(u)) \quad &- \quad \int_0^T\int_\Omega \operatorname{div}(f(u)F)f(u) + \operatorname{div}(h(u)F)h(u) - \int_0^T\int_\Omega \Delta g(u)(f(u)-h(u)) \\ &= \quad \int_0^T\int_\Omega f(u)(f(c)\,\overline{s} - f(u)\,\underline{s}) + h(u)(h(c)\,\overline{s} - h(u)\,\underline{s}).\end{aligned}$$

The right hand side of this equation is clearly bounded. The first term in the left side is also bounded (consider for example a primitive of $f(u)-h(u)$). So if we are able to bound the second term, by integrating by parts in the third term and using the fact that $\nabla f(u) = -\nabla h(u)$, we obtain (6.3.39)

Remark 6.3.6 *If we use the fact that $f'(u) \geq \alpha$ we have*

$$\int_0^T \int_\Omega \nabla g(u) \cdot \nabla u \leq C$$

which is an $L^2(0, T, H^1(\Omega))$ estimate for any function ζ satisfying $\zeta' = \sqrt{g'}$.

Let us now deal with the second term of the left hand side, *i.e.* the term concerning F. By summing equations 6.1.1 and 6.1.2, we already know that

$$\mathrm{div}(F) = \overline{s} - \underline{s}$$

so $\mathrm{div}(F)$ is bounded in $L^2(\Omega \times (0,T))$. By an easy calculation,

$$\mathrm{div}(f(u)F)f(u) = \frac{1}{2}\mathrm{div}(f(u)^2 F) + \frac{1}{2} f(u)^2 \mathrm{div} F,$$
$$\mathrm{div}(h(u)F)h(u) = \frac{1}{2}\mathrm{div}(h(u)^2 F) + \frac{1}{2} h(u)^2 \mathrm{div} F$$

and since $F \cdot n = 0$ on $\partial\Omega$,

$$\int_Q \mathrm{div}(f(u)^2 F) = \int_Q \mathrm{div}(h(u)^2 F) = 0.$$

Hence third term of the left hand side of is easily bounded and the proof is completed.

Step 2. The discrete counterpart : proof of (6.3.38).
In the following proof, we denote by C_i various real values which only depend on k_1, k_2, p_c, Ω, T, u_0, $\overline{s}$, $\underline{s}$, and not on D. Let us multiply (6.2.16) by $\delta t^n f(U_K^{n+1})$ and (6.2.17) by $\delta t^n h(U_K^{n+1})$ and sum the two equations thus obtained. Next we sum the result over $K \in \mathcal{T}$ and $n \in [\![0, N]\!]$. We obtain

$$\begin{aligned}
&\sum_{n=0}^{N} \sum_{K \in \mathcal{T}} |K|(U_K^{n+1} - U_K^n)(f(U_K^{n+1}) - h(U_K^{n+1})) \\
&-\sum_{n=0}^{N} \delta t^n \sum_{K \in \mathcal{T}} f(U_K^{n+1}) \sum_{L \in \mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} \delta_{K,L}^{n+1}(P) \\
&-\sum_{n=0}^{N} \delta t^n \sum_{K \in \mathcal{T}} h(U_K^{n+1}) \sum_{L \in \mathcal{N}_K} T_{K|L} k_{2,K|L}^{n+1} \delta_{K,L}^{n+1}(Q) \quad \leq C_8 \,.
\end{aligned}$$

Adding (6.2.16) and (6.2.17) gives

$$\sum_{L \in \mathcal{N}_K} T_{K|L} F_{K,L}^{n+1} = |K|(\overline{s}_K^{n+1} - \underline{s}_K^{n+1}). \tag{6.3.40}$$

where $F_{K,L}^{n+1}$ is the discrete counterpart of the total flux F, that is :

$$\begin{aligned}
F_{K,L}^{n+1} &= -k_{1,K|L}^{n+1} \delta_{K,L}^{n+1}(P) - k_{2,K|L}^{n+1} \delta_{K,L}^{n+1}(Q) \\
&= -(k_{1,K|L}^{n+1} + k_{2,K|L}^{n+1}) \delta_{K,L}^{n+1}(P) - k_{2,K|L}^{n+1} \delta_{K,L}^{n+1}(p_c(U)) \\
&= -(k_{1,K|L}^{n+1} + k_{2,K|L}^{n+1}) \delta_{K,L}^{n+1}(Q) + k_{1,K|L}^{n+1} \delta_{K,L}^{n+1}(p_c(U))
\end{aligned}$$

The first step of the estimate follows the continuous case, we introduce the total velocity flux F and the function $g(u)$ by writing $k_1\nabla P$ as a function of F and $\nabla p_c(u)$.
In the discrete case, the values of $U^{n+1}_{1,K|L}$ and $U^{n+1}_{2,K|L}$ can differ. Hence we shall need to decompose the numerical fluxes $k^{n+1}_{1,K|L}\delta^{n+1}_{K,L}(P)$ and $k^{n+1}_{2,K|L}\delta^{n+1}_{K,L}(Q)$ in the following way :

$$\begin{aligned}
&k^{n+1}_{1,K|L}\delta^{n+1}_{K,L}(P) = -f(U^{n+1}_{1,K|L})F^{n+1}_{K,L} + \Phi^{n+1}_{1,K,L} + R^{n+1}_{1,K,L},\\
&k^{n+1}_{2,K|L}\delta^{n+1}_{K,L}(Q) = -h(U^{n+1}_{2,K|L})F^{n+1}_{K,L} - \Phi^{n+1}_{2,K,L} - R^{n+1}_{2,K,L},\\
\text{with}&\\
&\Phi^{n+1}_{1,K,L} = -f(U^{n+1}_{1,K|L})k^{n+1}_{2,K|L}\delta^{n+1}_{K,L}(p_c(U)),\\
&\Phi^{n+1}_{2,K,L} = -h(U^{n+1}_{2,K|L})k^{n+1}_{1,K|L}\delta^{n+1}_{K,L}(p_c(U)),\\
\text{and}&\\
&R^{n+1}_{1,K,L} = f(U^{n+1}_{1,K|L})[k_2(U^{n+1}_{1,K|L}) - k_2(U^{n+1}_{2,K|L})]\delta^{n+1}_{K,L}(P)\\
&R^{n+1}_{2,K,L} = h(U^{n+1}_{2,K|L})[k_1(U^{n+1}_{1,K|L}) - k_1(U^{n+1}_{2,K|L})]\delta^{n+1}_{K,L}(Q)
\end{aligned}$$

Remark 6.3.7 *$\Phi^{n+1}_{1,K,L}$ and $\Phi^{n+1}_{2,K,L}$ are not very different of $\delta^{n+1}_{K,L}(g(U))$.*

In order to deal with time derivative terms, we once more use the inequality $(b-a)g'(b) \geq G(b) - G(a)$ for convex functions G such that $G' = f - h$ which yields that

$$-\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}|K|(U^{n+1}_K - U^n_K)(f(U^{n+1}_K) - h(U^{n+1}_K)) \leq \sum_{K\in\mathcal{T}}|K|(G(U^{N+1}_K) - G(U^0_K)) \leq C_9\,. \quad (6.3.41)$$

Gathering by edges, and remarking that $\delta^{n+1}_{K,L}(f(U)) + \delta^{n+1}_{K,L}(h(U)) = 0$ (this is a direct consequence of $f + h = 1$), we obtain then

$$\begin{aligned}
&\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}f(U^{n+1}_K)\sum_{L\in\mathcal{N}_K}T_{K|L}f(U^{n+1}_{1,K|L})F^{n+1}_{K,L} \quad +\\
&\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}h(U^{n+1}_K)\sum_{L\in\mathcal{N}_K}T_{K|L}h(U^{n+1}_{2,K|L})F^{n+1}_{K,L} \quad +\\
\frac{1}{2}&\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}T_{K|L}(\Phi^{n+1}_{1,K,L} + \Phi^{n+1}_{2,K,L} + R^{n+1}_{1,L,K} + R^{n+1}_{2,L,K})\delta^{n+1}_{K,L}(f(U)) \qquad \leq C_9\,,
\end{aligned} \quad (6.3.42)$$

Since $f + h = 1$, multiplying (6.3.40) by $f(U^{n+1}_K) + h(U^{n+1}_K)$, summing over $K \in \mathcal{T}$ and substracting from (6.3.42) yields:

$$\begin{aligned}
&\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}f(U^{n+1}_K)\sum_{L\in\mathcal{N}_K}T_{K|L}(f(U^{n+1}_{1,K|L}) - f(U^{n+1}_K))F^{n+1}_{K,L} \quad +\\
&\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}h(U^{n+1}_K)\sum_{L\in\mathcal{N}_K}T_{K|L}(h(U^{n+1}_{2,K|L}) - h(U^{n+1}_K))F^{n+1}_{K,L} \quad +\\
\frac{1}{2}&\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}T_{K|L}(\Phi^{n+1}_{1,K,L} + \Phi^{n+1}_{2,K,L} + R^{n+1}_{1,L,K} + R^{n+1}_{2,L,K})\delta^{n+1}_{K,L}(f(U)) \qquad \leq C_{10}\,,
\end{aligned} \quad (6.3.43)$$

Using the equality $b(a-b) = -\frac{1}{2}(a-b)^2 + \frac{1}{2}(a^2-b^2)$, we get from (6.3.43),

$$\begin{aligned}
&-\tfrac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} \sum_{\substack{L\in\mathcal{N}_K\\(K,L)\notin\mathcal{E}_1^{n+1}}} T_{K|L}(f(U_K^{n+1})-f(U_L^{n+1}))^2\ F_{K,L}^{n+1}\\
&+\tfrac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} \sum_{\substack{L\in\mathcal{N}_K\\(K,L)\notin\mathcal{E}_1^{n+1}}} T_{K|L}(f^2(U_K^{n+1})-f^2(U_L^{n+1}))\ F_{K,L}^{n+1}\\
&-\tfrac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} \sum_{\substack{L\in\mathcal{N}_K\\(K,L)\notin\mathcal{E}_2^{n+1}}} T_{K|L}(h(U_K^{n+1})-h(U_L^{n+1}))^2\ F_{K,L}^{n+1} \qquad (6.3.44)\\
&+\tfrac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} \sum_{\substack{L\in\mathcal{N}_K\\(K,L)\notin\mathcal{E}_2^{n+1}}} T_{K|L}(h^2(U_K^{n+1})-h^2(U_L^{n+1}))\ F_{K,L}^{n+1}\\
&+\tfrac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} T_{K|L}(\Phi_{1,K,L}^{n+1}+\Phi_{2,K,L}^{n+1}+R_{1,L,K}^{n+1}+R_{2,L,K}^{n+1})\delta_{K,L}^{n+1}(f(U)) \qquad \le C_{10}\,.
\end{aligned}$$

If we denote by T_2 and T_4 the second and the fourth term in (6.3.44), we have

$$T_2 \;=\; \frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} (f(U_K^{n+1}))^2 \Big(\sum_{\substack{L\in\mathcal{N}_K\\(K,L)\notin\mathcal{E}_1^{n+1}}} T_{K|L}F_{K,L}^{n+1} - \sum_{\substack{L\in\mathcal{N}_K\\(L,K)\notin\mathcal{E}_1^{n+1}}} T_{K|L}F_{L,K}^{n+1}\Big)$$

But $(L,K)\notin\mathcal{E}_1^{n+1} \Leftrightarrow (K,L)\in\mathcal{E}_1^{n+1}$ and $F_{K,L} = -F_{L,K}$, so

$$\begin{aligned}
T_2 \;&=\; \frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{K\in\mathcal{T}} (f(U_K^{n+1}))^2 \sum_{L\in\mathcal{N}_K} T_{K|L}F_{K,L}^{n+1}\\
&\le\; C_{11}
\end{aligned}$$

and in the same way $T_4 \le C_{12}$.
Therefore if we develop all the terms, we obtain

$$\begin{aligned}
&\frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{(K,L)\in\mathcal{E}_1^{n+1},(K,L)\in\mathcal{E}_2^{n+1}} T_{K|L}|\delta_{K,L}^{n+1}(f(U))|^2(F_{K,L}^{n+1}+F_{K,L}^{n+1})\\
&+\frac{1}{2}\sum_{n=0}^{N}\delta t^n \sum_{(K,L)\in\mathcal{E}_1^{n+1},(K,L)\notin\mathcal{E}_2^{n+1}} T_{K|L}|\delta_{K,L}^{n+1}(f(U))|^2(F_{K,L}^{n+1}+F_{L,K}^{n+1})\\
&+\sum_{n=0}^{N}\delta t^n \sum_{(K,L)\in\mathcal{E}_1^{n+1}} T_{K|L}(\Phi_{1,K,L}^{n+1}+\Phi_{2,K,L}^{n+1}+R_{1,L,K}^{n+1}+R_{2,L,K}^{n+1})\delta_{K,L}^{n+1}(f(U)) \qquad \le C_{13}\,. \qquad (6.3.45)
\end{aligned}$$

Since $F_{K,L}^{n+1}+F_{L,K}^{n+1} = 0$, (6.3.45) leads to

$$\sum_{n=0}^{N} \delta t^n \sum_{(K,L)\in\mathcal{E}_1^{n+1}} T_{K|L} D_{K,L}^{n+1} \leq C_{14}$$

where

$$\begin{aligned} &D_{K,L}^{n+1} = |\delta_{K,L}^{n+1}(f(U))|^2 F_{K,L}^{n+1} + (\Phi_{1,K,L}^{n+1} + \Phi_{2,K,L}^{n+1} + R_{1,L,K}^{n+1} + R_{2,L,K}^{n+1})\delta_{K,L}^{n+1}(f(U)),\\ &\forall (K,L) \in \mathcal{E}_1^{n+1}, (K,L) \in \mathcal{E}_2^{n+1} \end{aligned} \tag{6.3.46}$$

and

$$\begin{aligned} &D_{K,L}^{n+1} = (\Phi_{1,K,L}^{n+1} + \Phi_{2,K,L}^{n+1} + R_{1,L,K}^{n+1} + R_{2,L,K}^{n+1})\delta_{K,L}^{n+1}(f(U)),\\ &\forall (K,L) \in \mathcal{E}_1^{n+1}, (K,L) \notin \mathcal{E}_2^{n+1}. \end{aligned} \tag{6.3.47}$$

Let us first study $D_{K,L}^{n+1}$ for $(K,L) \in \mathcal{E}_1^{n+1} \cap \mathcal{E}_2^{n+1}$. Since $U_{1,K|L}^{n+1} = U_{2,K|L}^{n+1} = U_K^{n+1}$, it is clear that $R_{1,L,K}^{n+1} = R_{2,L,K}^{n+1} = 0$ and we have

$$D_{K,L}^{n+1} = \quad \delta_{K,L}^{n+1}(f(U))(\delta_{K,L}^{n+1}(f(U))F_{K,L}^{n+1} - \quad 2\frac{k_1(U_K^{n+1})k_2(U_K^{n+1})}{k_1(U_K^{n+1}) + k_2(U_K^{n+1})}\delta_{K,L}^{n+1}(p_c(U))).$$

If we assume that $U_K^{n+1} \leq U_L^{n+1}$, then $\delta_{K,L}^{n+1}(p_c(U)) \leq 0$ and $F_{K,L}^{n+1} \geq -k_2(U_K^{n+1})(\delta_{K,L}^{n+1}(p_c(U)))$, which leads to

$$\begin{aligned} D_{K,L}^{n+1} &\geq -[\delta_{K,L}^{n+1}(f(U))k_2(U_K^{n+1}) + 2\frac{k_1(U_K^{n+1})k_2(U_K^{n+1})}{k_1(U_K^{n+1}) + k_2(U_K^{n+1})}]\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(p_c(U))\\ &\geq -[f(U_K^{n+1}) + f(U_L^{n+1})]k_2(U_K^{n+1})\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(p_c(U))\\ &\geq -\frac{k_1(U_L^{n+1})k_2(U_K^{n+1})}{k_1(U_L^{n+1}) + k_2(U_L^{n+1})}\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(p_c(U))\\ &\geq \alpha\delta_{K,L}^{n+1}(f(U))\int_{U_K^{n+1}}^{U_L^{n+1}} \frac{k_1(a)k_2(a)}{k_1(a)+k_2(a)}(-p_c{}'(a))da.\\ &\geq C_{15}\,\delta_{K,L}^{n+1}(g(U)) \end{aligned} \tag{6.3.48}$$

Remark 6.3.8 *To deduce the third inequality from the second one, we used the fact* $f(U_K^{n+1}) \geq 0$.

We get the same estimate in the case $U_K^{n+1} \geq U_L^{n+1}$ is similar, it only suffices by symmetry to exchange k_1 by k_2 and f by h.

We now study $D_{K,L}^{n+1}$ for $(K,L) \in \mathcal{E}_1^{n+1}$, $(K,L) \notin \mathcal{E}_2^{n+1}$. We have $\delta_{K,L}^{n+1}(P) \leq 0$ and $\delta_{K,L}^{n+1}(Q) \geq 0$, which yields $\delta_{K,L}^{n+1}(p_c(U)) \geq 0$, and therefore $U_K^{n+1} \geq U_L^{n+1}$.

$$\begin{aligned} D_{K,L}^{n+1} \quad = -\delta_{K,L}^{n+1}(f(U)) \quad &[f(U_K^{n+1})k_2(U_L^{n+1})\delta_{K,L}^{n+1}(p_c(U)) + \delta_{K,L}^{n+1}(k_2(U))\delta_{K,L}^{n+1}(P)\\ &+h(U_L^{n+1})(k_1(U_L^{n+1})\delta_{K,L}^{n+1}(p_c(U)) + \delta_{K,L}^{n+1}(k_1(U))\delta_{K,L}^{n+1}(Q)] \end{aligned}$$

Now we use the symmetry of the problem in p and q. We can express $\delta_{K,L}^{n+1}(Q)$ in function of $\delta_{K,L}^{n+1}(P)$ or the contrary. In the first case we obtain

$$\begin{aligned} D_{K,L}^{n+1} &= -[(f(U_K^{n+1})k_2(U_L^{n+1}) + h(U_L^{n+1})(k_1(U_L^{n+1}))]\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(p_c(U)) \\ &\quad -[f(U_K^{n+1})\delta_{K,L}^{n+1}(k_2(U)) + h(U_L^{n+1})\delta_{K,L}^{n+1}(k_1(U))]\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(P). \end{aligned}$$

In the second case, we obtain

$$\begin{aligned} D_{K,L}^{n+1} &= -[(f(U_K^{n+1})k_2(U_K^{n+1}) + h(U_L^{n+1})(k_1(U_K^{n+1}))]\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(p_c(U)) \\ &\quad -[f(U_K^{n+1})\delta_{K,L}^{n+1}(k_2(U)) + h(U_L^{n+1})\delta_{K,L}^{n+1}(k_1(U))]\delta_{K,L}^{n+1}(f(U))\delta_{K,L}^{n+1}(Q). \end{aligned}$$

Since $\delta_{K,L}^{n+1}(P) \leq 0$ and $\delta_{K,L}^{n+1}(Q) \geq 0$, one of the two terms,

$$[f(U_K^{n+1})\delta_{K,L}^{n+1}(k_2(U)) + h(U_L^{n+1})\delta_{K,L}^{n+1}(k_1(U))]\delta_{K,L}^{n+1}(P)$$

or

$$[f(U_K^{n+1})\delta_{K,L}^{n+1}(k_2(U)) + h(U_L^{n+1})\delta_{K,L}^{n+1}(k_1(U))]\delta_{K,L}^{n+1}(Q)$$

is non negative. Moreover, one has

$$f(U_K^{n+1})k_2(U_K^{n+1}) + h(U_L^{n+1})k_1(U_K^{n+1}) \geq \frac{k_1(U_K^{n+1})k_2(U_L^{n+1})}{k_1(U_K^{n+1}) + k_2(U_L^{n+1})}$$

and

$$f(U_K^{n+1})k_2(U_L^{n+1}) + h(U_L^{n+1})k_1(U_L^{n+1}) \geq \frac{k_1(U_K^{n+1})k_2(U_L^{n+1})}{k_1(U_K^{n+1}) + k_2(U_L^{n+1})},$$

one gets

$$\begin{aligned} D_{K,L}^{n+1} &\geq \alpha\delta_{K,L}^{n+1}(f(U)) \int_{U_K^{n+1}}^{U_L^{n+1}} \frac{k_1(a)k_2(a)}{k_1(a)+k_2(a)}(-p_c{}'(a))da \\ &\geq C_{16}\, \delta_{K,L}^{n+1}(g(U))\delta_{K,L}^{n+1}(f(U)) \end{aligned} \tag{6.3.49}$$

Using (6.3.47), (6.3.48) and (6.3.49) yield (6.3.38). □

6.4 Compactness properties

Following [EGH00b] one may deduce from (6.3.33) the following property:

Corollary 6.4.1 (Pressure space translates) *Let C_1 be given by Proposition 6.3.3, then for any $\xi \in \mathbb{R}^d$, the following inequalities holds:*

$$\int_0^T \int_{\Omega_\xi} [p_{\mathcal{D}}(x+\xi,t) + p_g(u_{\mathcal{D}})(x+\xi,t) - p_{\mathcal{D}}(x,t) - p_g(u_{\mathcal{D}})(x,t)]^2 dxdt \leq C_1\,|\xi|(2size(\mathcal{T}) + |\xi|) \tag{6.4.50}$$

Similarly, we easily deduce from the estimate on $g(u)$ given in Proposition 6.3.5 the following property :

Corollary 6.4.2 **($g(u)$ Space translates)** *Assume that (H) are fulfilled and let C_7 be given by Proposition 6.3.5, then for any $\xi \in \mathbb{R}^d$ such that $|\xi| \leq \mathrm{diam}(\Omega)$, the following inequality holds*

$$\int_0^T \int_{\Omega_\xi} [g(u_{\mathcal{D}}(x+\xi,t)) - g(u_{\mathcal{D}}(x,t))]^2 dxdt \leq C_7\, L_g |\xi| (2size(\mathcal{T}) + |\xi|), \tag{6.4.51}$$

where $\Omega_\xi = \{x \in \mathbb{R}^d, [x, x+\xi] \subset \Omega\}$.

In the proof of convergence below, an important argument is the strong compactness of the sequence $g(u_{D_n})$ in $L^2(\Omega \times (0,T))$.We already have an estimate of the space translates, we also need an estimate on the time translates of $g(u_D)$ to apply Kolmogorov's theorem. This estimate is given in the following proposition.

Proposition 6.4.1 **($g(u)$ Time translates)** *Under hypotheses (H), one denotes by g the function defined, for all $s \in [0,1]$, by $g(s) = -\int_0^s \frac{k_1(a)k_2(a)}{k_1(a)+k_2(a)}\, {p_c}'(a)da$. Let D be a finite volume discretization of $\Omega \times (0,T)$ in the sense of Definition 6.2.3 and let (U_D, P_D) be a solution of the finite volume scheme (6.2.15)-(6.2.19) .*
Then there exists C_{17} , which only depends on k_1, k_2, p_c, Ω, T, u_0, $\overline{s}$, $\underline{s}$, and not on D, such that, for all $\tau \in (0,T)$, the following discrete estimate hold

$$\int_0^{T-\tau} \int_\Omega [g(u_{\mathcal{D}}(x,t+\tau)) - g(u_{\mathcal{D}}(x,t))]^2\, dx\, dt \leq C_{17}\, \tau \tag{6.4.52}$$

Proof. The proof of this estimate is close to the one of the parabolic case (Cf [EGH00b])
Step 1. Proof in the continuous case . We give here the analogue of this proof in the continuous case. The main argument is that we have a bound on $k_1(u)(\nabla p)^2$ and on $(\nabla g(u))^2$ in $L^1(\Omega \times (0,T))$ and that we know u_t by using the equation. By using Fubini-Tonelli theorem, we have

$$\int_\Omega \int_0^{T-\tau} [g(u(x,t+\tau)) - g(u(x,t))]^2\, dt\, dx \quad = \quad \int_0^{T-\tau} A(t)dt$$

Where $A(t) = \int_\Omega [g(u(x,t+\tau) - g(u(x,t))]^2\, dt\, dx$.
Since g is a Lipschitz function with Lipschitz constant Lg,

$$\begin{aligned} A(t) \quad &\leq \quad L_g \int_\Omega (g(u(x,t+\tau)) - g(u(y,t)))(u(x,t+\tau) - u(x,t))dx \\ &\leq \quad L_g \int_\Omega (g(u(x,t+\tau)) - g(u(x,t))) \int_t^{t+\tau} u_t(x,\theta)d\theta\, dx \\ &\leq \quad L_g \int_\Omega \int_t^{t+\tau} (g(u(x,t+\tau)) - g(u(x,t)))[\mathrm{div}(k_1(u)\nabla p)(x,\theta) \\ &\qquad\qquad + f(c)\,\overline{s}(x,\theta) - f(u(x,\theta))\,\underline{s}(x,\theta)]\, d\theta\, dx \end{aligned}$$

Now if we develop and make an integrate by parts in x, we obtain

$$\begin{aligned}
A(t) \quad \leq \quad & L_g\int_\Omega \int_t^{t+\tau} k_1(u(x,\theta))\nabla p(x,\theta)\nabla g(u(x,t+\tau))\,d\theta\,dx \\
& -L_g\int_\Omega \int_t^{t+\tau} k_1(u(x,\theta))\nabla p(x,\theta)\nabla g(u(x,t))\,d\theta\,dx \\
& +L_g\int_\Omega \int_t^{t+\tau} (f(c)\,\overline{s}(x,\theta) - f(u(x,\theta))\,\underline{s}(x,\theta))(g(u(x,t+\tau)) - g(u(x,t)))\,d\theta\,dx
\end{aligned}$$

Thanks to Young Inequality, we get

$$A(t) \leq \frac{Lg}{2}(2A_1(t) + A_2(t) + A_3(t) + 2A_4(t))$$

with

$$\begin{aligned}
A_1(t) &= \int_t^{t+\tau}\int_\Omega k_1(u(x,\theta))(\nabla p(x,\theta))^2\,d\theta dx \\
A_2(t) &= \int_t^{t+\tau}\int_\Omega k_1(u(x,\theta))(\nabla g(u(t)))^2\,d\theta\,dx \\
A_3(t) &= \int_t^{t+\tau}\int_\Omega k_1(u(x,\theta))(\nabla g(u(t+\tau)))^2\,d\theta\,dx \\
A_4(t) &= C(g)\int_t^{t+\tau}\int_\Omega \overline{s}(x,\theta) + \underline{s}(x,\theta)\,d\theta\,dx
\end{aligned}$$

Then using Fubini theorem, and the bound obtained in the preceding propositions, we can say that

$$\int_0^T A(t) \leq C\tau$$

<u>*Step 2. Proof in the continuous case* .</u> For $t \in [0,t)$, let us denote by $n(t)$ the integer $n \in [\![0, N+1]\!]$ such that $t \in [t^n, t^{n+1})$. We can write

$$\int_0^{T-\tau}\int_\Omega (g(u_{\mathcal{D}}(x,t+\tau)) - g(u_{\mathcal{D}}(x,t)))^2 dxdt = \int_0^{T-\tau} A(t)dt,$$

with, for a.e. $t \in (0, T-\tau)$,

$$A(t) = \sum_{K\in\mathcal{T}} |K|(g(u_K^{n(t+\tau)+1}) - g(u_K^{n(t)+1}))^2.$$

Since g is non decreasing and Lipschitz continuous with constant L_g, one gets

$$\begin{aligned}
A(t) \quad \leq \quad & L_g\sum_{K\in\mathcal{T}} |K|(g(u_K^{n(t+\tau)+1}) - g(u_K^{n(t)+1}))(u_K^{n(t+\tau)+1} - u_K^{n(t)+1}) \\
\leq \quad & L_g\sum_{K\in\mathcal{T}} (g(u_K^{n(t+\tau)+1}) - g(u_K^{n(t)+1}))\sum_{n=n(t)+1}^{n(t+\tau)} |K|(U_K^{n+1} - U_K^n) \\
\leq \quad & L_g\sum_{K\in\mathcal{T}} (g(u_K^{n(t+\tau)+1}) - g(u_K^{n(t)+1}))\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \begin{pmatrix} \displaystyle\sum_{L\in\mathcal{N}_K} T_{K|L}k_{1,K|L}^{n+1}\delta_{K,L}^{n+1}(P) + \\ |K|(f(c_K^{n+1})\,\overline{s}_K^{n+1} - f(U_K^{n+1})\,\underline{s}_K^{n+1}) \end{pmatrix}.
\end{aligned}$$

Gathering by edges, we get

$$
\begin{aligned}
A(t) \quad \leq \quad & L_g \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}} \left(\sum_{L\in\mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} \delta_{K,L}^{n+1}(P)(g(u_K^{n(t+\tau)+1}) - g(u_L^{n(t+\tau)+1})) \right) \\
- \quad & L_g \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}} \left(\sum_{L\in\mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} (\delta_{K,L}^{n+1}(P)(g(u_K^{n(t)+1}) - g(u_L^{n(t)+1})) \right) \\
+ \quad & L_g \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}} \Big(|K|(f(c_K^{n+1})\, \overline{s}_K^{n+1} - f(U_K^{n+1})\, \underline{s}_K^{n+1})(g(u_K^{n(t+\tau)+1}) - g(u_K^{n(t)+1}) \Big).
\end{aligned}
$$

Thanks to the Young inequality, we get

$$
A(t) \quad \leq \quad \frac{L_g}{2}(2A_1(t) + A_2(t) + A_3(t) + 2A_4(t))
$$

with

$$
\begin{aligned}
A_1(t) \quad &= \quad \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}} \left(\sum_{L\in\mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} |\delta_{K,L}^{n+1}(P)|^2 \right) \\
A_2(t) \quad &= \quad \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}} \left(\sum_{L\in\mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} (g(u_K^{n(t+\tau)+1}) - g(u_L^{n(t+\tau)+1}))^2 \right) \\
A_3(t) \quad &= \quad \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}} \left(\sum_{L\in\mathcal{N}_K} T_{K|L} k_{1,K|L}^{n+1} (g(u_K^{n(t)+1}) - g(u_L^{n(t)+1}))^2 \right) \\
A_4(t) \quad &= \quad \sum_{n=n(t)+1}^{n(t+\tau)} C_{18} \sum_{K\in\mathcal{T}} (\, \overline{s}_K^{n+1} + \underline{s}_K^{n+1}).
\end{aligned}
$$

Using the following lemma, the proof of which is given at section 5.3.1 the proof of (6.4.52) easily follows.

Lemma 6.4.1 *Let $T > 0$, $\tau \in (0,T)$, $k \in (0,T)$ and $(a^n)_{n\in\mathbb{N}}$ be a family of non negative real values. Then*

$$
\int_0^{T-\tau} \sum_{n=n(t)+1}^{n(t+\tau)} a^{n+1} dt \leq \tau \sum_{n=0}^{N} a^{n+1},
$$

and for any $\zeta \in [0,\tau]$

$$
\int_0^{T-\tau} \sum_{n=n(t)+1}^{n(t+\tau)} a^{n(t+\zeta)+1} dt \leq \tau \sum_{n=0}^{N} a^{n+1}.
$$

□

We are now able to prove that up to a subsequence, the scheme is convergent. By using Propositions 6.3.23, 6.4.2 and 6.4.1, we find that $g(u_{D_m})$ satisfies the hypothesis of the Kolmogorov's theorem with $Q = \Omega \times (0,T)$ if $\lim_{n\to+\infty} size(D_m) = 0$ so there exists a function $\tilde{g} \in L^2(0,T;H^1(\Omega))$ such that up to a subsequence, $g(u_{D_m}) \to \tilde{g}$ in $L^2(\Omega \times (0,T))$. Since g is strictly increasing and u_D remains bounded, it follows that $u_{D_m} \to u := g^{-1}(\tilde{g})$ in $L^2(\Omega \times (0,T))$.

Therefore $p_g(u_{D_m}) \to p_g(u)$ in $L^2(\Omega \times (0,T))$.

Using the discrete Poincaré inequality, we get that $P_{D_m} + p_g(u_{D_m})$ remains bounded in $L^2(\Omega \times (0,T))$, and therefore there exists $\tilde{p} \in L^2(0,T;H^1(\Omega))$ such that $P_{D_m} + p_g(u_{D_m}) \rightharpoonup \tilde{p}$ weakly in $L^2(\Omega \times (0,T))$. It follows that $P_{D_m} \rightharpoonup p := \tilde{p} - p_g(u)$. The last step in the demonstration of the convergence theorem 6.2.1 it to prove that (u,p) is a solution. This is the aim of the next section.

6.5 Study of the limit

Proposition 6.5.1 *Assume hypothesis (H) are fulfilled and let D_m be a sequence of time-space discretizations of $\Omega \times (0,T)$ such that $\lim_{n\to+\infty} size(D_m) = 0$. Assume moreover that the sequence of corresponding approximate solutions (u_{D_m}, p_{D_m}) converges to (u,p) as $n \to +\infty$. Then (u,p) is a weak solution of (6.1.1)-(6.1.7) in the sense of Definition 6.1.1.*

The proof of this theorem follows classical guidelines. Let $\varphi \in \mathcal{D}(\Omega \times (0,T))$. We multiply Equations (6.2.16) and (6.2.17) by $\varphi(x_K, t^{n+1})$ and sum over $K \in \mathcal{T}$ and $n \in [\![0,N]\!]$. Then there remains to show that the discrete terms converge to the corresponding integrals terms. The only difficulty concerns the integral term $\int_0^T \int_\Omega k_1(u)\nabla p \cdot \nabla\varphi$ and $\int_0^T \int_\Omega k_2(u)\nabla q \cdot \nabla\varphi$. The proof of Theorem 6.5.1 is then a direct consequence of the following lemmas.

Lemma 6.5.1 (Weak-Strong convergence) *Under hypothese (H), let D_m be a sequence of time-space discretizations of $\Omega \times (0,T)$ such that $lim_{m\to\infty} size(D_m) = 0$. Let $k \in L^2(\Omega \times (0,T))$. Let $K_{D_m} \in X_{D_m}$, for $m \in \mathbb{N}$, such that the corresponding discrete approximation k_{D_m} converges to k in $L^2(\Omega \times (0,T))$. Let $V_m \in X_{D_m}$, for $m \in \mathbb{N}$, such that*

$$\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} T_{K|L}(V_L^{n+1} - V_K^{n+1})^2 \leq C_{19}$$

and that there exists $v \in L^2(0,T;H^1(\Omega))$ with $V_{D_m} \rightharpoonup v$ weakly in $L^2(\Omega \times (0,T))$.
Let φ be a test function. Let T_m be defined by

$$T_m = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}} \varphi(x_K, t^{n+1}) \sum_{L\in\mathcal{N}_K} T_{K|L} k_{K,L}^{n+1}(V_L^{n+1} - V_K^{n+1}),$$

where, for all $(K,L) \in \mathcal{E}$, $k_{K,L}^{n+1} \in \{k_K^{n+1}, k_L^{n+1}\}$ and $k_{K,L}^{n+1} = k_{L,K}^{n+1}$.
Then

$$\lim_{m\to+\infty} T_m = \int_0^T \int_\Omega k(x,t)\nabla v(x,t)\nabla\varphi(x,t)dxdt.$$

Proof. see Chapter 5. $\square$

Lemma 6.5.2 *Let us assume that u_{D_m} converges to u in $L^2(\Omega \times (0,T))$ with*

$$\sum_{n=0}^{N} \delta t^n \sum_{K \in \mathcal{T}} \sum_{L \in \mathcal{N}_K} T_{K|L} \delta_{K,L}^{n+1} (g(U))^2 \leq C_{20}$$

Let φ be a test function. Let T_m be defined by

$$T_m = -\sum_{n=0}^{N} \delta t^n \sum_{K \in \mathcal{T}} \varphi(x_K, t^{n+1}) \sum_{L \in \mathcal{N}_K} T_{K|L} k_1(U_{K,L}^{n+1}) \delta_{K,L}^{n+1}(p_g(U)),$$

where, for all $(K,L) \in \mathcal{E}$, $U_{K,L}^{n+1} \in \{U_K^{n+1}, U_L^{n+1}\}$ and $U_{K,L}^{n+1} = U_{L,K}^{n+1}$.
Then

$$\lim_{m \to +\infty} T_m = \int_0^T \int_\Omega \nabla g(u)(x,t) \nabla \varphi(x,t) dx dt.$$

Proof. Gathering by edges, the term T_m can be rewritten as :

$$T_m = \frac{1}{2} \sum_{n=0}^{N} \delta t^n \sum_{(K,L) \in \mathcal{E}} T_{K|L} k_1(U_{K,L}^{n+1}) \delta_{K,L}^{n+1}(p_g(U))(\varphi(x_L, t^{n+1}) - \varphi(x_K, t^{n+1})).$$

Let

$$T_{1m} = \frac{1}{2} \sum_{n=0}^{N} \delta t^n \sum_{(K,L) \in \mathcal{E}} T_{K|L} \delta_{K,L}^{n+1}(g(U))(\varphi(x_L, t^{n+1}) - \varphi(x_K, t^{n+1})).$$

Using Lemma 6.5.1, one has :

$$\lim_{m \to +\infty} T_{1m} = \int_0^T \int_\Omega \nabla g(u)(x,t) \nabla \varphi(x,t) dx dt.$$

We now study $T_{1m} - T_m$. Thanks to the regularity of Φ and the monotonicity properties of k_1, k_2 and p_c, one has :

$$|T_{1m} - T_m| \leq C_{21} \sum_{n=0}^{N} \delta t^n \sum_{(K,L) \in \mathcal{E}} T_{K|L} d_{K|L} A_{K,L}^{n+1},$$

with

$$A_{K,L}^{n+1} = |k_1(U_K^{n+1}) - k_1(U_L^{n+1})|\,|p_g(U_K^{n+1}) - p_g(U_L^{n+1})|.$$

In the non degenerate case, it is straightforward that $|T_{1m} - T_m|$ tends to 0 as m tends to infinity. Indeed, if $k_1 \geq k_{1,min}$, then $|p_g(U_K^{n+1}) - p_g(U_L^{n+1})| \leq \frac{1}{k_{1,min}}|g(U_K^{n+1}) - g(U_L^{n+1})|$ and thus we get

$$A_{K,L}^{n+1} \leq \frac{Lip_{k_1}}{k_{1,min}}\delta_{K,L}^{n+1}(U)\delta_{K,L}^{n+1}(g(U))$$

Therefore thanks to Estimate (6.3.38) on $g(u)$, we get $|T_{1m} - T_m| \leq Csize(D_m)$ which clearly tends to zero.

In the degenerate case, under Hypotheses (H), the proof is somewhat more complicated. For a given set of values $(\varepsilon_{K|L}^{n+1})_{K|L\in\mathcal{E}}$, we separate the edges where $|\delta_{K,L}^{n+1}(U)|$ is less than $\varepsilon_{K|L}^{n+1}$ from the edges where $|\delta_{K,L}^{n+1}(U)|$ is larger than $\varepsilon_{K|L}^{n+1}$, we then choose $\varepsilon_{K|L}^{n+1}$ to optimize the bound. Let us give the computations:

- if $|U_K^{n+1} - U_L^{n+1}| \leq \varepsilon_{K|L}^{n+1}$, then thanks to the the regularity of k_1 and the Hölder continuity of p_g (consequence (C4) of Hypotheses (H)) one gets

$$A_{K,L}^{n+1} \leq Lip_{k_1}C_{1-\beta_2,p_g}(\varepsilon_{K|L}^{n+1})^{2-\beta_2}$$

- if $|U_K^{n+1} - U_L^{n+1}| \geq \varepsilon_{K|L}^{n+1}$, thanks to Inequality 6.1.10 (take $a = U_K^{n+1}$ and $b = U_L^{n+1}$), one has

$$A_{K,L}^{n+1} \leq C(\alpha,\beta_1,\beta_2)|\delta_{K,L}^{n+1}(k_1(U))|\frac{|\delta_{K,L}^{n+1}(g(U))|}{|\delta_{K,L}^{n+1}(U)|} \leq C(\alpha,\beta_1,\beta_2)\frac{1}{\alpha\varepsilon_{K|L}^{n+1}}\delta_{K,L}^{n+1}(U)\delta_{K,L}^{n+1}(g(U))$$

Thus

$$T_{K|L}d_{K|L}A_{K,L}^{n+1} \leq C(\alpha,\beta_1,\beta_2,Lip_{k_1},C_{1-\beta_2,p_g})[T_{K|L}\delta_{K,L}^{n+1}(U)\delta_{K,L}^{n+1}(g(U))\frac{d_{K|L}}{\varepsilon_{K|L}^{n+1}} + m(K|L)d_{K|L}\frac{(\varepsilon_{K|L}^{n+1})^{2-\beta_2}}{d_{K|L}}]$$

We know that $\displaystyle\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}T_{K|L}\delta_{K,L}^{n+1}(U)\delta_{K,L}^{n+1}(g(U))$ and $\displaystyle\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}_K}m(K|L)d_{K|L}$ are bounded, so to have a good bound on $A_{K,L}^{n+1}$, we can take $\varepsilon_{K|L}^{n+1}$ such that $\frac{d_{K|L}}{\varepsilon_{K|L}^{n+1}} = \frac{(\varepsilon_{K|L}^{n+1})^{2-\beta_2}}{d_{K|L}}$, that is

$$\varepsilon_{K|L}^{n+1} = d_{K|L}^{\frac{2}{3-\beta}}$$

Then $\frac{d_{K|L}}{\varepsilon_{K|L}^{n+1}} \leq size(\mathcal{T})^{\frac{1-\beta}{3-\beta}}$ and

$$|T_{1m} - T_m| \leq Csize(\mathcal{T})^{\frac{1-\beta_2}{3-\beta_2}}$$

which proves Lemma 6.5.2, since $\beta_2 < 1$.
□

6.6 Preuve de l'inégalité (6.1.10)

Let us prove that for all $(a, b) \in [0, 1]^2$,

$$(b-a)\int_a^b \frac{k_2(s)}{k_1(s)+k_2(s)}\,(-p_c{}'(s))ds \leq C(\alpha,\beta_1,\beta_2)|\int_a^b \frac{k_1(s)k_2(s)}{k_1(s)+k_2(s)}\,(-p_c{}'(s))ds$$

Proof.
We can suppose $0 \leq a \leq b \leq 1$. First of all, by using the hypothesis, we easily have:

$$\frac{(b-a)\int_a^b \frac{k_2(s)}{k_1(s)+k_2(s)}\,(-p_c{}'(s))ds}{\int_a^b \frac{k_1(s)k_2(s)}{k_1(s)+k_2(s)}\,(-p_c{}'(s))ds} \quad \leq \quad C(\alpha,\beta_1,\beta_2)\frac{(b-a)\int_a^b s^{-\beta_1}(1-s)^{1-\beta_2}ds}{\int_a^b s^{1-\beta_1}(1-s)^{1-\beta_2}ds}$$

The right term is a continuous function of a and b on the subset $\{b-a>0\}$ then it is bounded on the compact subset $\{b-a \geq \frac{1}{2}\}$ and there only remains to deal for example with the subsets $\{a \geq \frac{1}{4}\}$ and $\{b \leq \frac{3}{4}\}$ of the triangle $\{0 \leq a \leq b \leq 1\}$.

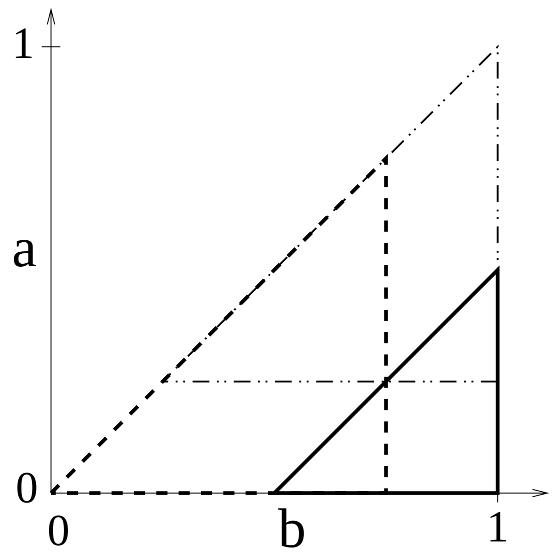


Fig 14. *The covering of* $\{0 \leq a \leq b \leq 1\}$ *by the subsets* $\{b-a \geq \frac{1}{2}\}$, $\{a \geq \frac{1}{4}\}$ *and* $\{b \leq \frac{3}{4}\}$.

Suppose that $b \leq \frac{3}{4}$. Then $(\frac{1}{4})^{1-\beta_2} \leq (1-s)^{1-\beta_2} \leq 1$. Then we can include this term in the function $C(\alpha,\beta_1,\beta_2)$ and easily obtain a bound by integrating the functions $s^{1-\beta_1}$ and $s^{-\beta_1}$. It remains to show that

$$\frac{(b-a)(b^{2-\beta_1}-a^{2-\beta_1})}{b^{1-\beta_1}-a^{1-\beta_1}}$$

is bounded. If $a = 0$ this is equal to 1. If $a \neq 0$, we can factorize by a. Then if $x = \frac{b}{a}$, we only need to prove that $\rho(x) = \frac{(x-1)(x^{2-\beta_1}-1)}{x^{1-\beta_1}-1}$ is bounded on $(1, +\infty)$. This function is continuous, it tends to 1 if $x \to \infty$ and it tends to 0 when $x \to 1$. So, we have finished with the case $b \leq \frac{3}{4}$.
Now, if $a \geq \frac{1}{4}$ there is no difficulty since $(\frac{1}{4})^{1-\beta_1} \leq s^{-\beta_1}$ and $s^{-\beta^1} \leq (\frac{1}{4})^{-\beta_1}$ and $(b-a) \leq \frac{3}{4}$. The proof is complete. □

Chapitre 7

Résultats numériques

7.1 Comparaison des deux schémas en 1D

L'implémentation des deux schémas en une dimension d'espace ne pose pas de difficulté technique. Les résulats numériques présentés ici sont obtenus par le schéma totalement implicite avec le logiciel de calcul scientifique Matlab 5.

7.1.1 Test comparatif (système découplé)

Dans un premier temps, nous avons fait des tests sur un cas très simple afin de comparer les deux schémas. Le domaine physique est l'intervalle $\Omega = (0, 1)$ et les paramètres sont définis de la manière suivante :

$$k_1(x) = x,\ k_2(x) = 1 - x,\ \text{ et }\ p_c(x) = 1 - x. \tag{7.1}$$

Dans ce cas, comme $M(x) = k_1(x) + k_2(x) = 1$, le système est découplé, et on peut facilement calculer le champ de pression à partir des sources, puisque l'équation elliptique sur la pression globale est simplement l'équation différentielle linéaire du deuxième ordre à coefficients constants suivante :

$$-(M(u)p')' = \overline{s} - \underline{s}. \tag{7.2}$$

où

$$\overline{s} = 40\,\mathbb{1}_{[0.1,0.3]} + 20\,\mathbb{1}_{[0.7,0.9]}$$

et

$$\underline{s} = 60\,\mathbb{1}_{[0.4,0.6]}$$

Les fonctions f et φ peuvent être facilement calculées à partir des paramètres et on obtient :

$$f(x) = x,\ \text{ et }\ \varphi(x) = x - \frac{x^2}{2}. \tag{7.3}$$

On choisit une condition initiale avec une discontinuité en $x = 0.5$: $u_0(x) = 0.9$ si $x < 0.5$, $u_0(x) = 0.3$ si $x > 0.5$. Pour finir, on fixe la concentration en injection : $c = 0.8$.
Comme on s'y attendait, les deux schémas donnent de bons résultats, le principe du maximum est respecté et l'algorithme de Newton converge rapidement, comme le montre la figure ci dessous.

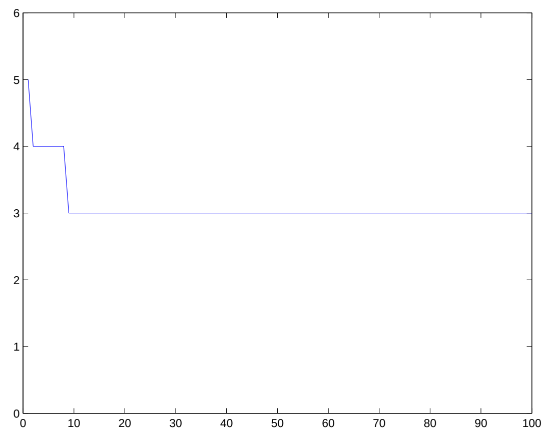


Fig 15. *Nombre d'itérations de la méthode de Newton en fonction de n pour le schéma des "mathématiciens". Test 1.*

Les résultats des deux schémas sont quasiment identiques, la différence entre les deux solutions est très faible.

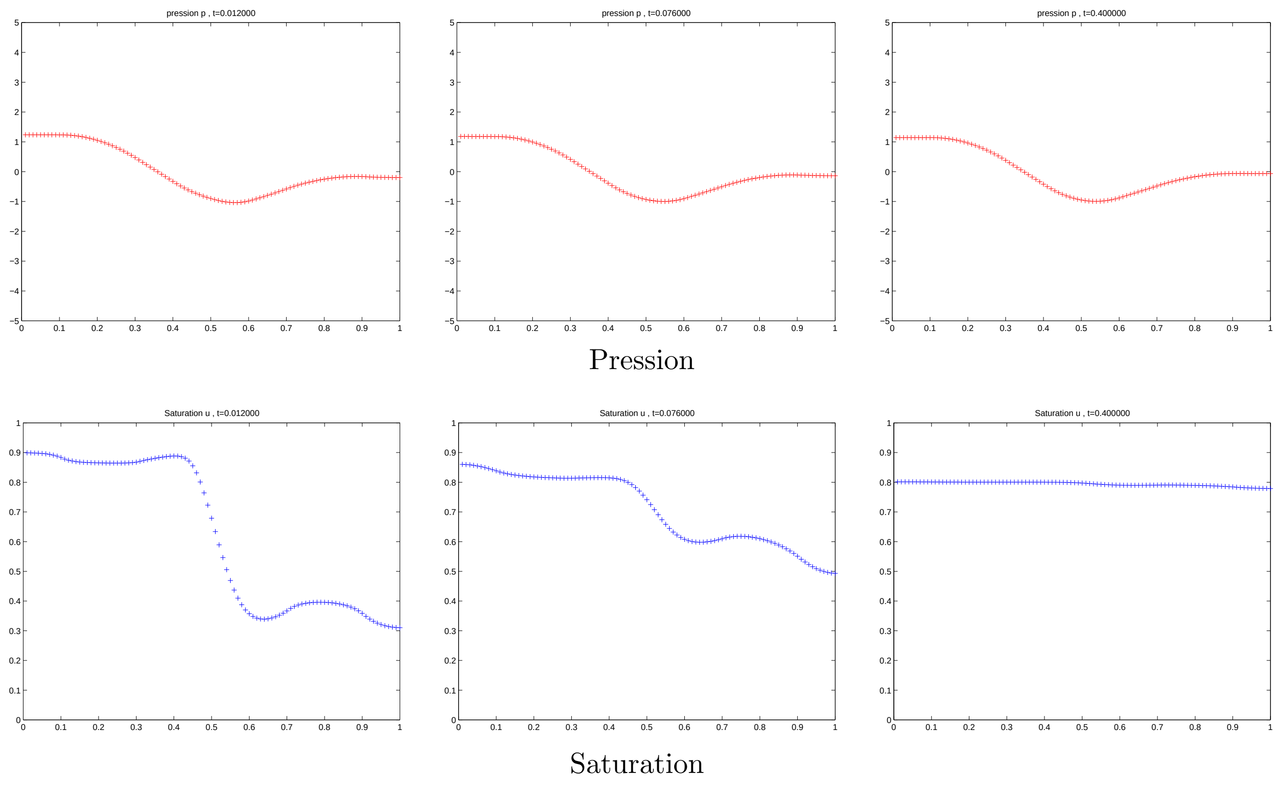


Fig 16. *Solutions obtenues par le schéma des "mathématiciens". Test 1.*

Pour la saturation, ce sont visiblement les sources qui font directement évoluer la solution. La discontinuité disparaît dans les premiers pas de temps. Pour des temps grands, la saturation évolue vers un équilibre qui est la concentration en injection, $c = 0.8$.

7.1.2 Test en absence de pression capillaire

Pour mieux comprendre le rôle de la pression capilaire, nous avons fait le même test avec $p_c = 0$ cette fois ci.

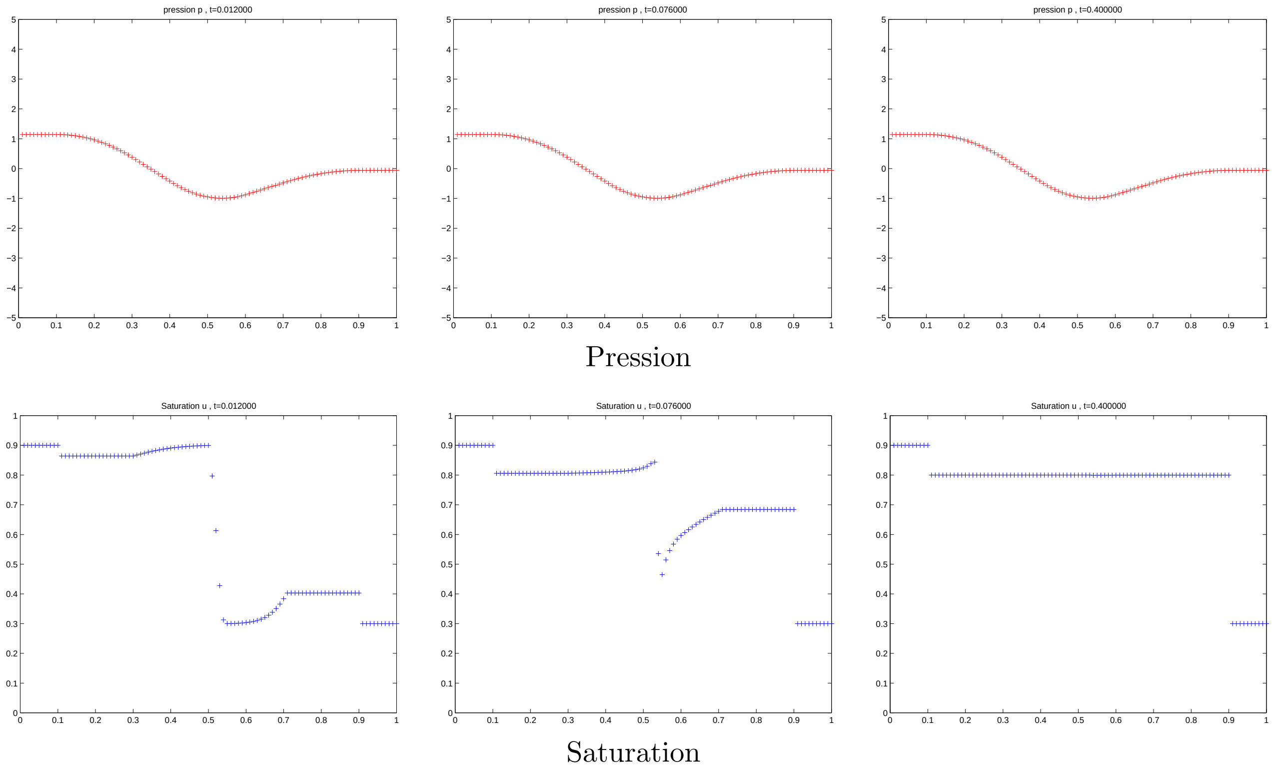


Fig 17. *Solutions obtenues par le schéma des "pétroliers". Test 2.*

Le comportement de la solution n'est plus du tout le même. L'équation sur la saturation étant ici une équation hyperbolique et non plus une équation parabolique, les discontinuités ne disparaissent pas et au contraire de nouvelles discontinuités apparaissent. Le régime stationnaire atteint n'est plus le même. En effet, le champ de pression est constant aux extrémités du domaine puisque les lignes de courant vont des sources vers le puits, et les sources s'annulent sur $[0; 0.1]$ et $[0.9; 1]$. Ainsi, la saturation reste constante au cours du temps aux deux extrémités, ce qui était impossible en présence de pression capillaire. On voit aussi que le gradient de pression a peu d'effet en comparaison de l'afflux des sources, tant que la concentration du fluide est éloignée de la concentration d'injection. Cette interprétation peut se comprendre en regardant l'équation sur la saturation qui se réécrit de la manière suivante losrqu'on developpe le terme de convection :

$$u_t = [f(c) - f(u)]\,\overline{s} + f'(u)\nabla u \cdot M(u)\nabla p. \tag{7.4}$$

7.1.3 Test avec couplage des équations

Le couplage des équation est automatique dès que M n'est plus constante. Ce couplage est d'autant plus fort que M a de fortes variations sur la plage des saturations admises. En pratique, la variation de M n'est cependant pas importante. Nous avons tout de même fait un test pour observer les différences avec le cas précédemment étudié. Pour cela, nous avons pris les fonctions suivantes :

$$k_1(x) = x^3,\ k_2(x) = (1-x)^3, \text{ et } p_c(x) = 1 - x.$$

Pour ne pas compliquer la résolution, nous regardons uniquement cette fois-ci les résulats obtenus avec le schéma des pétroliers 1D. Ce qui apparait d'abord lors des tests numériques est la plus grande difficulté à résoudre le sytème non linéaire (qui cette fois est vraiment couplé) qui s'ajoute à la necessité de réduire le pas de temps , comme le montre la figure ci dessous.

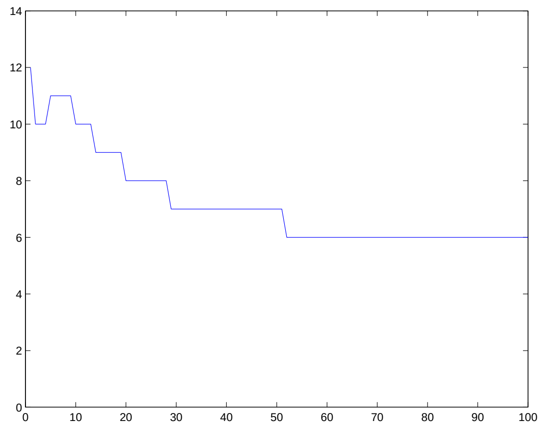
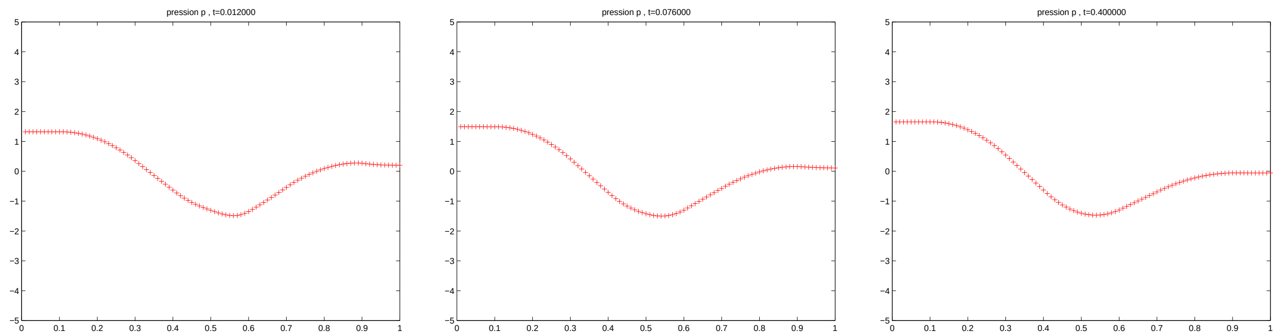
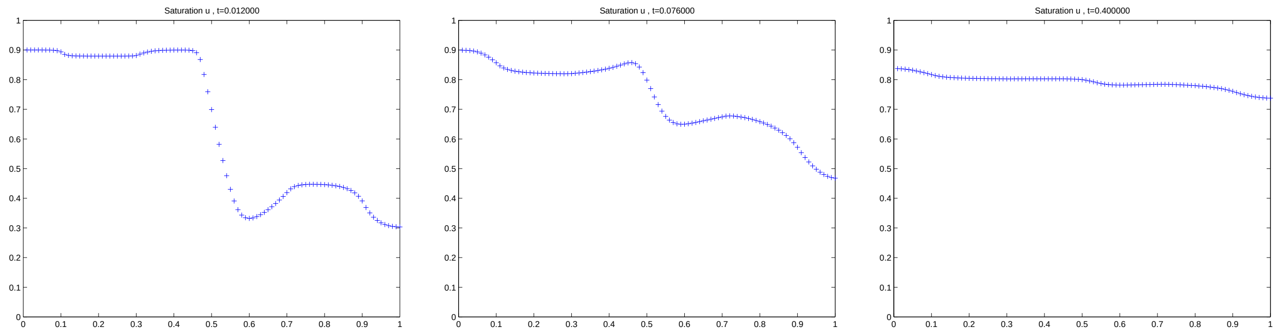


Fig 18. *Itérations de la méthode de Newton en fonction de n pour le schéma des "pétroliers". Test 3.*

En ce qui concerne l'évolution de la saturation, les différences sont difficiles à voir. Ce qui confirme que les sources sont le principal vecteur de l'évolution dans l'équation parabolique sur la saturation. Les résultats obtenus font apparaître globalement les mêmes tendances que pour le Test 1.



Pression



Saturation

Fig 19. *Solutions obtenues par le schéma des "pétroliers". Test 3.*

7.2 Calcul en 2D par le schéma “des pétroliers”

La programmation du schéma des mathématiciens sur maillage 2D triangulaire ne nous a pas permis d’obtenir de solution par cette méthode. Le domaine utilisé avec un maillage rectangulaire était dans notre programme limité au domaine carré $\Omega = [0,1]^2$. Les résulats obtenus sur le cas test (cf [Ohl97], [AA99]) sont donnés en Section 5.4. La comparaison des deux schémas n’a pas apporté plus d’enseignements sur maillage rectangulaire qu’en 1D. Toutefois, nous avons testé le schéma des pétroliers dans plusieurs situations qui pouvaient poser des problèmes afin d’éprouver sa robustesse.

Nous avons fait le même test que celui fait sur maillage rectangulaire avec le schéma des mathématiciens au paragraphe 5.4. Nous avons utilisé le maillage 2D sur le carré construit par raffinement. Comme on peut le voir sur les figures ci dessous, la solution obtenue est très proche des résultats obtenus avec le “schéma des mathématiciens”.

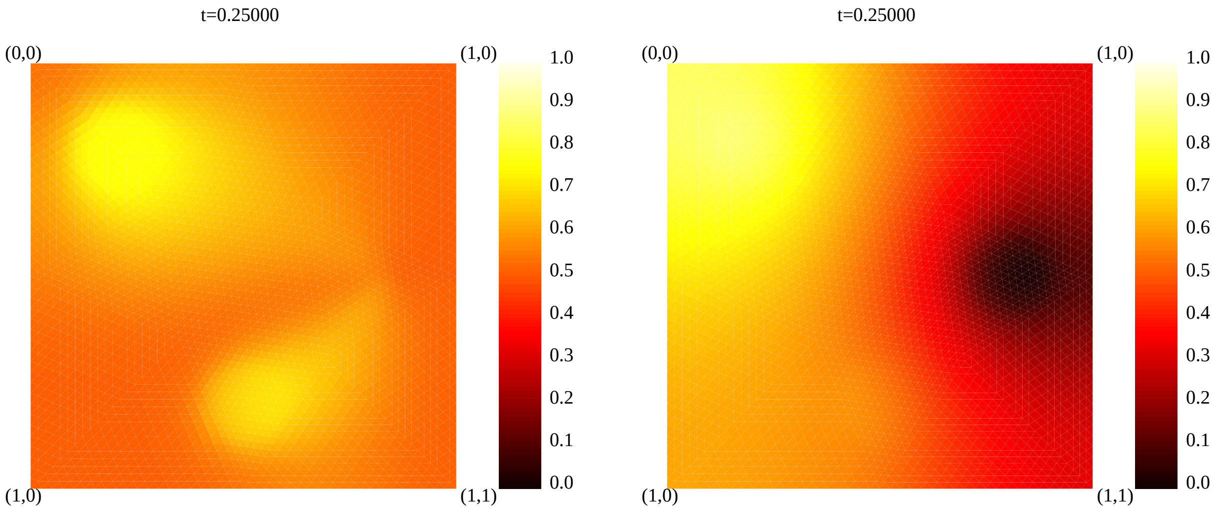


Fig 20. *Saturation et Pression à la date* $t = 0.25$

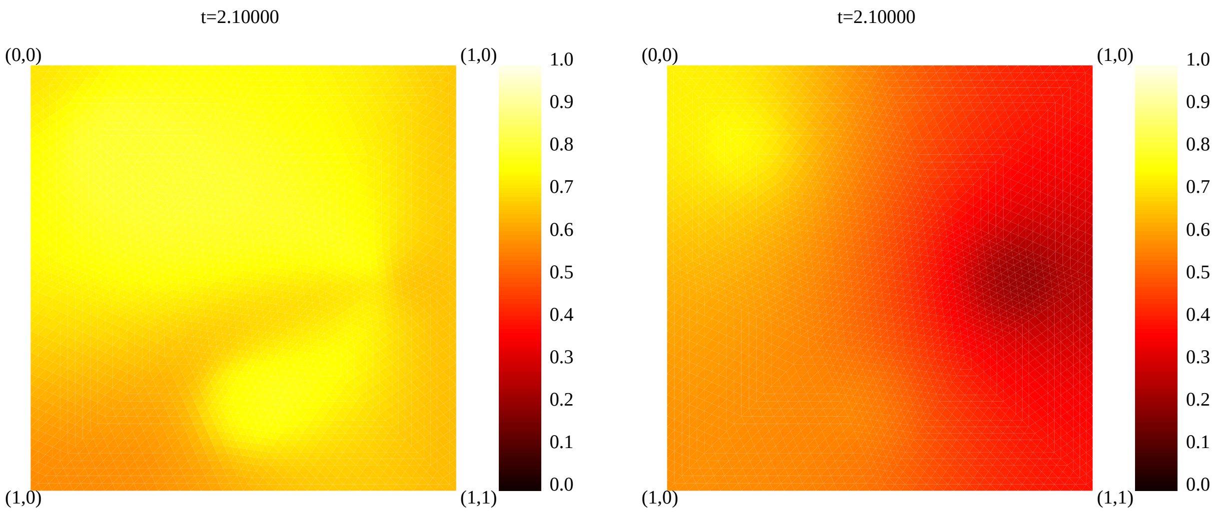


Fig 21. *Saturation et Pression à la date* $t = 2.1$

Conclusion

Résultats

Dans la première partie de ce mémoire, nous avons étudié la convergence de schémas volumes finis pour les équations paraboliques dégénérées dans le cas où le terme de diffusion est faiblement dégénéré ou fortement dégénéré avec respectivement des conditions au bord de type Neumann homogène et de type Dirichlet non homogènes. Il ressort principalement de cette étude les conclusions suivantes :

- La solution approchée obtenue par un schéma volumes finis sur des maillages admissibles vérifie plusieurs inégalités discrètes, qui sont des estimations a priori rigoureuses correspondant respectivement à une estimation sur la norme de u dans $L^\infty(\Omega)$ et à une estimation sur la norme de $\varphi(u)$ dans $L^2(0,T,H^1(\Omega))$. Ces deux estimations permettent d'obtenir une propriété de compacité faible de l'ensemble des solutions approchées notée $u_{\mathcal{T},\delta t}$ ou u_D dans $L^\infty(\Omega)$ et de compacité forte de $\varphi(u_{\mathcal{T},\delta t})$ ou $\varphi(u_D)$ dans $L^2(0,T,L^2(\Omega))$, cette dernière propriété de compacité découlant directement du théorème de Kolmogorov.

- Néanmoins, chacun des cas traités est différent pour les raisons suivantes :

  – Dans le cas faiblement dégénéré (Chapitre 1), comme φ est strictement croissante on déduit directement de la compacité forte de $\varphi(u_{\mathcal{T},\delta t})$ une compacité forte sur $u_{\mathcal{T}}$ qui suffit pour démontrer que les valeurs d'adhérences sont des solutions faibles du problème considéré. Cette convergence est suffisante car dans le cas faiblement dégénéré, il y a unicité d'une solution faible.

  – Dans les cas où le terme de diffusion est fortement dégénéré (Chapitres 2 et 3), φ est simplement croissante au sens large, donc on ne peut rien déduire a priori sur u_D à partir de la connaissance de $\varphi(u_D)$. Ce manque de compacité peut alors être compensé comme dans l'étude des équations hyperboliques (cf [EGGH98][CCL95][Vil94]) en utilisant les solutions de type mesures d'Young [DiP85]. D'autre part, la notion de solution faible n'étant plus suffisante pour assurer l'unicité, il est nécessaire de faire appel à une formulation entropique.

    De l'étude du cas fortement dégénéré avec conditions au bord de type Dirichlet homogènes et $\mathbf{q}\cdot\mathbf{n} = 0$ (Chapitre 2), il ressort que l'adaptation de la preuve d'unicité pour les équations hyperboliques de SN Kruzkov aux équations paraboliques nécessite de prendre en compte séparemment les zones ou φ est strictement croissante. Une bonne manière de tenir compte de ces zones est de formuler les inégalités obtenues avec les entropies convexes dont la dérivée est du type $\eta'(\varphi(u))$, avec η convexe, qui sont par nature plus régulières puisque $\varphi(u)$ vérifie des estimations plus fortes que la solution u elle même. A la manière de José Carrillo [Car99], on peut alors construire une preuve d'unicité pour les solutions processus entropiques.

    Le cas fortement dégénéré avec conditions au bord de type Dirichlet homogène sans condition sur $\mathbf{q}\cdot\mathbf{n}$ (Chapitre 3) n'entre pas dans le cadre de l'étude faite au chapitre précédent. Les travaux publiés depuis le début des années 70 sur les conditions au bord pour les équations hyperboliques dans le cadre BV (cf [BlRN79], [CCL95]) puis plus récemment dans le cadre L^∞ (cf [Ott96a], [Vov00]) ont montré qu'il était nécessaire de formuler également les conditions au bord en un sens entropique. En utilisant les travaux les plus récents, qui sont beaucoup mieux adaptés aux méthodes volumes finis, nous avons donné une formulation intégrale avec des entropies simples (demi-entropies de Kruzkov), qui étend les formulations de José Carrillo [Car99] et de Julien Vovelle [Vov00]. Les hypothèses sur la condition au bord, nécessaires pour démontrer la convergence du schéma, sont identiques à celles utilisées pour le cas $\mathbf{q} \cdot \mathbf{n} = 0$ au chapitre 2. La preuve d'unicité est basée sur l'idée que l'on peut remplacer dans les inégalités d'entropie le réel κ par un relèvement encore noté $\bar{u}$ de la condition au bord. Il ressort de ce travail que la méthode de dédoublement des variables de Kruzkov peut être adaptée pour montrer l'unicité des solutions entropiques pour les équations paraboliques hyperboliques avec données au bord non homogènes et que les schémas volumes finis sont bien adaptés à ce type de problème.

La deuxième partie de ce mémoire a permis de mettre en évidence les résultats suivants, concernant les méthodes volumes finis pour les écoulements diphasiques en milieu poreux :

- Le schéma volumes finis "des mathématiciens", qui consiste à réécrire le système sous la forme d'une équation elliptique sur la pression couplée avec une équation parabolique hyperbolique sur la saturation par l'utilisation de la pression globale de Chavent [CJ86] est convergent. La preuve peut se faire par l'utilisation des méthodes déjà connues dans le cadre des équations elliptiques et des équations paraboliques. Elle est valable dans le cas où le terme de diffusion est faiblement dégénéré.

- Le schéma volumes finis des "pétroliers", qui consiste à discrétiser directement le système des équations de conservation en effectuant un décentrage phase par phase (Chapitre 6), est lui aussi convergent. La démonstration de la convergence du schéma fait appel à des estimations a priori originales que la présence simultanée de deux décentrages différents rend ardues. Il ressort de cette étude que la symétrie du problème par rapport aux deux phases est essentiel à la preuve de la convergence du schéma et qu'elle joue un rôle important dans la robustesse du schéma.

- Les applications numériques ne permettent pas de comparer les deux schémas du point de vue des résultats obtenus, même lorsque les deux méthodes numériques peuvent être testées (1D ou 2D rectangle) Néanmoins, des faiblesses au niveau de la robustesse du schéma des "mathématiciens" sont apparues. En revanche, aucune difficulté n'apparaît lors de la résolution du système associé au schéma des pétroliers sur un maillage triangulaire. Par ailleurs, il est indéniable qu'au-delà des performances, la programmation du schéma des pétroliers est beaucoup plus proche de la physique du problème, donc plus simple à mettre en oeuvre. Enfin, contrairement au schéma des "mathématiciens" ce schéma à l'avantage de pouvoir s'étendre au cas compressible et compositionnel.

Perspectives

Les résultats mis en évidence dans la première partie du mémoire laissent entrevoir une première perspective qui est l'amélioration des hypothèses sur la donnée au bord dans le cas général étudié au chapitre 3. Au-delà du souci d'optimisation des hypothèses, il est également intéressant d'envisager une étude concernant les estimations d'erreur. Dans les trois premiers chapitres de ce mémoire, la convergence du schéma est démontrée par des estimations a priori et l'utilisation de fonctions tests régulières, donc sans évaluation de l'erreur entre solution exacte et solution approchée. En ce qui concerne l'estimation d'erreur, j'entreprends un travail avec Raphaèle Herbin et Robert Eymard sur le schéma volumes finis exposé au chapitre 2, en s'appuyant sur les résultats obtenus dans le cadre de l'approximation visqueuse [EGH00a].

La deuxième partie du mémoire a mis en évidence la robustesse du schéma des pétroliers, mais aussi la difficulté de la preuve de convergence. La première perspective serait d'envisager des mobilités relatives plus générales comprenant des fonctions quadratiques ou cubiques par exemple, les mobilités réelles étant plus faciles à représenter à l'aide de telles fonctions. Enfin, le schéma des "pétroliers" étant un schéma industriel, une seconde perspective qui s'offre naturellement est d'étendre la preuve faite au chapitre 6 à des milieux non isotropes et non homogènes, tout en tenant compte des termes gravitaires.

Annexe A

Unicité des solutions faibles pour les équations paraboliques faiblement dégénérées par dualité

Dans cette annexe, on démontre l'unicité de la solution u du problème

$$u_t - \Delta\varphi(u) + \operatorname{div}(\mathbf{q}f(u)) = 0.$$

La démonstration s'inspire de la preuve faite dans [EGH00b] pour l'équation

$$u_t - \Delta\varphi(u) = 0$$

D'autres preuves sont possibles, mais le principe de démonstration par dualité est intéressant en soi. Il faut noter qu'on utilise fortement les estimations sur les solutions faibles, par exemple le fait que les solutions soient bornées est primordial. Le spécificité de la preuve par dualité est d'utiliser les théorèmes d'existence dans le cas linéaire et les estimations associées pour montrer l'unicité d'un problème totalement nonlinéaire.

A.1 Définition de la notion de solution

On étudie ici des solutions au sens faible du problème et la formulation intégrale contient la condition initiale et les conditions limites au bord.

Definition A.1.1 (Solution faible) *On appelle solution faible du problème considéré toute fonction* $u \in L^\infty(\Omega \times (0,T))$ *vérifiant*

$\forall \psi \in \mathcal{C}$,

$$\int_0^T\int_\Omega u(x,t)\psi_t(x,t) + f(u(x,t))\mathbf{q}(x,t).\nabla\psi(x,t) + \varphi(u(x,t))\Delta\psi(x,t)\,dxdt \\ + \int_\Omega u_0(x,t)\psi(x,0)\,dx = 0,$$

où $\mathcal{C} = \{\psi \in C^{2,1}(\bar{\Omega} \times [0,T])$ *telles que* $\psi(.,T) = 0$ *et* $\nabla\psi.\mathbf{n} = 0\}$.

A.2 Unicité par la méthode duale

Le principe de cette méthode est très simple à comprendre et s'adapte a d'autres problèmes nonlinéaires du même type.

Soient u_1 et u_2 deux solutions faibles. Si l'on note u_d la différence entre ces deux fonctions alors, pour toute fonction $\psi \in \mathcal{C}$, on a

$$\int_0^T \int_\Omega u_d(x,t)[\psi_t(x,t) + \mathbf{q}(x,t)F(x,t).\nabla\psi(x,t) + \Phi(x,t)\Delta\psi(x,t)]dxdt{=}0, \tag{A.1}$$

où $F(x,t) = \tilde{f}(u_1(x,t), u_2(x,t))$ et $\Phi(x,t) = \tilde{\varphi}(u_1(x,t), u_2(x,t))$, avec pour toute fonction g dérivable $\tilde{g}(a,b) = \frac{g(a)-g(b)}{a-b}$ si $a \neq b$ et $g'(a)$ dans le cas contraire.

Definition A.2.1 (Problème dual) *Le problème dual associé à u_1, u_2 et à une fonction χ est*

$$\psi \in C^{2,1}(\bar{\Omega} \times [0,T]) \text{ et } \begin{cases} \psi_t + \mathbf{q}F.\nabla\psi + \Phi\Delta\psi = \chi \\ \nabla\psi.\mathbf{n} = 0 \\ \psi(T) = 0 \end{cases}$$

A.2.1 Principe de la démonstration d'unicité

Supposons que l'on sache démontrer l'existence d'une fonction ψ solution du problème dual pour toute fonction χ assez régulière (par exemple dans $\mathcal{C}_c^\infty(\Omega \times (0,T))$. D'après (A.1), on aurait

$$\int_0^T \int_\Omega u_d(x,t)\chi(x,t)dxdt = 0, \forall \chi \in \mathcal{C}_c^\infty(\Omega \times (0,T))\,).$$

Par un argument classique de densité, il serait alors facile de démontrer que $u_d = 0$ au sens $L^\infty(\Omega \times (0,T))$.

Malheureusement, les seules hypothèses $f \in \mathcal{C}^1(\mathbb{R})$ et $\varphi \in \mathcal{C}^1(\mathbb{R})$ strictement croissante sont insuffisantes pour espérer montrer l'existence d'une solution au problème dual.
En fait, la principale difficulté provient du fait que le problème initial est dégénéré, c'est à dire que φ' peut s'annuler, donc Φ *n'est pas minorée par une constante strictement positive.*

A.2.2 Théorème d'existence pour un problème proche

On considère dans un premier temps le problème dual dans le cas ou F, v et Φ ont de bonnes propriétés. Le premier théorème suivant est admis :

Theorem A.2.1 (Existence pour le problème dual régularisé) *Si on suppose que F , $\mathbf{q}$ et Φ sont de classe $\mathcal{C}^\infty$ sur $\bar{\Omega}\times[0,T]$, et qu'il existe $\delta > 0$ tel que $\Phi(x,t) \geq \delta$, alors pour toute fonction $\chi \in \mathcal{C}_c^\infty(\Omega\times(0,T))$, le problème dual admet au moins une solution.*

Les hypothèses de ce théorème sont un cas particulier des hypothèses du théorème 5.3 dans [LSU67].

A.2.3 Principe du maximum et comparaison

Dans cette partie, nous allons montrer que si la fonction χ est comprise dans l'intervalle [m,M], alors $\psi(x,t)$ est compris entre $m(T-t)$ et $M(T-t)$.

Les théories classiques étant le plus souvent écrites avec une donnée initiale prise au temps $t=0$ pour une évolution suivant les $t>0$, nous effectuons un changement de variable $s=T-t$ qui nous ramène au problème :

$$\psi \in C^{2,1}(\bar{\Omega} \times [0,T]) \text{ et } \begin{cases} \psi_t - \Phi\Delta\psi - \mathbf{q}F.\nabla\psi = \chi \\ \nabla\psi.\mathbf{n} = 0 \\ \psi(0) = 0 \end{cases}$$

Le résultat qui nous sera utile par la suite est le suivant :

Theorem A.2.2 (Principe de comparaison) *Soit θ une fonction vérifiant*

$$\theta \in C^{2,1}(\bar{\Omega} \times [0,T]) \textit{ et } \begin{cases} \theta_t - \Phi\Delta\theta - \mathbf{q}F.\nabla\theta \leq 0 \\ \nabla\theta.\mathbf{n} \leq 0 \\ \theta(0) \leq 0 \end{cases}$$

alors , $\theta(x,t) \leq 0 \quad \forall (x,t) \in \bar{\Omega} \times [0,T]$.

Remark A.2.1 *Si on change le sens des inégalités dans les hypothèses et la conclusion, ce théorème est bien sûr encore vrai, puisque on peut appliquer le théorème A.2.2 à $-\theta$.*

Ce théorème découle directement du principe du maximum pour des équations paraboliques. Les résultats utiles et leur démonstration figurent en Section A.3

En appliquant à la différence de deux fonctions le théorème A.2.2, on obtient un principe de comparaison de la solution au problème avec des "*sur-solutions*" et des "*sous-solutions*" de celui-ci. Ceci permet notamment de démontrer le résultat suivant :

Proposition A.2.1 *Si ψ est une solution du problème dual régularisé associé à χ, et si $m \leq \chi \leq M$, alors*

$$m(T-t) \leq \psi(x,t) \leq M(T-t), \quad \forall (x,t) \in \bar{\Omega} \times [0,T].$$

En particulier, ψ est bornée en norme $L^p(\Omega \times (0,T))$ pour tout p, $1 \leq p \leq \infty$, indépendemment de la borne infèrieure δ de Φ.

A.2.4 Une estimation sur $\Delta\psi$ et $\nabla\psi$

Pour comparer le problème dual et le problème dual régularisé, on aura besoin d' estimations concernant $\Delta\psi$ et $\nabla\psi$ en fonction des données, à savoir les coefficients Φ, $\mathbf{q}$ et F, et le second membre χ. Dans cette première estimation, nous utiliront bien sur le fait que $\Phi \geq \delta$, mais en dehors de cela, nous nous contenterons d'utiliser le fait que Φ, v et F sont des fonctions bornées et que les intégrations par parties se font sans problème, car les fonctions sont régulières.

Proposition A.2.2 (Estimation L^∞) *Soit ψ une solution au problème dual régularisé, et soit M_Φ , $M_{\mathbf{q}}$ et M_F des majorants de Φ, $|\mathbf{q}|$ et $|F|$. Il existe $C>0$ ne dépendant que de χ, M_Φ, $M_{\mathbf{q}}$, M_F, Ω et T telle que*

$$\|\Delta\psi\|_{L^2(\Omega\times(0,T))} \leq \frac{C}{\delta^2}$$

et

$$\|\nabla\psi\|_{L^2(\Omega\times(0,T))} \le \frac{C}{\delta}.$$

Démonstration .

On multiplie l'équation par $\Delta\psi$ et on intègre le résultat sur $\Omega\times(t,T)$. Après intégration par parties on obtient l'équation

$$\frac{1}{2}\int_\Omega |\nabla\psi(t)|^2 + \int_0^t\int_\Omega \Phi|\Delta(\psi)|^2 = \int_0^t\int_\Omega -\nabla\chi.\nabla\psi - \int_0^t\int_\Omega \Delta\psi F\mathbf{q}.\nabla\psi. \tag{A.2}$$

Le premier terme dans le membre de droite de (A.2) n'est pas génant, par contre, le second terme a le meme degré d'homogénéité par rapport à ψ que le terme de gauche. On ne peut pas l'éliminer directement. On utilise alors l'inégalité auxiliaire suivante

$$\|\nabla\psi\|^2_{L^2(\Omega\times(0,T))} \le \|\psi\|_{L^2(\Omega\times(0,T))}\|\Delta\psi\|_{L^2(\Omega\times(0,T))}. \tag{A.3}$$

qui découle directement de

$$\|\nabla\psi\|^2_{L^2(\Omega\times(0,T))} = \int_0^T\int_\Omega |\nabla\psi|^2 = -\int_0^T\int_\Omega \psi\Delta\psi$$

(A.2) prise à la date $t = T$ et (A.3) associés à l'estimation sur ψ obtenue par le principe de comparaison permettent d'aboutir à l'inégalité

$$\int_0^T\int_\Omega \Phi|\Delta(\psi)|^2 \le C_1\|\Delta\psi\|^{\frac{1}{2}}_{L^2(\Omega\times(0,T))} + C_2\|\Delta\psi\|^{\frac{3}{2}}_{L^2(\Omega\times(0,T))}, \tag{A.4}$$

où C_1 et C_2 sont du meme type que C dans la proposition.

Finalement, en utilisant des inégalités de Young, on obtient

$$\delta(1-\alpha)\|\Delta\psi\|_{L^2(\Omega\times(0,T))} \le \frac{C_1'(\alpha)}{\delta^2} + \frac{C_2'(\alpha)}{\delta^3},$$

où $\alpha < 1$ est un réel fixé.

L'estimation sur $\Delta\psi$ s'en déduit directement grace au fait que $\delta \le M_\Phi$, puis en utilisant (A.3), on obtient celle concernant $\nabla\psi$.

Remark A.2.2 *Si on essaie d'utiliser l'équation (A.2) pour obtenir une estimation de $\nabla\psi(t)$, alors on tombe sur une majoration du type*

$$\|\nabla\psi(t)\|_{L^2(\Omega)} \le \frac{C}{\delta^{\frac{3}{2}}}.$$

Par conséquent, on a une "moins bonne" estimation sur la norme $L^\infty(0,T,H^1(\Omega))$ que sur la norme $L^2(0,T,H^1(\Omega))$.

A.2.5 Résultat d'unicité et démonstration

Voici le résultat que nous allons démontrer

Theorem A.2.3 (Unicité) *si φ^{-1} est α-Hölder avec $\alpha \geq \frac{1}{2}$, alors le problème posé admet au plus une solution faible.*

Démonstration.

Le travail d'estimation préliminaire permet de démontrer de manière concise ce résultat.

Soit $\chi \in \mathcal{C}_c^\infty(\Omega \times (0,T))$.

Le théorème d'existence pour le problème dual nécessite entre autres que Φ soit minorée par un réel strictement positif. Il est donc inapplicable tel quel.

Soit $\delta > 0$ fixé arbitrairement ($\delta \leq M_\Phi$)

$\Phi_\delta = max(\delta, \Phi)$ est minorée par δ et est encore dans $L^{\infty(\Omega \times (0,T))}$. Cependant, on n'a pas d'hypothèses concernant la régularité de Φ_δ, F et $\mathbf{q}$. Qu'à cela ne tienne en effectuant une convolution par un noyau régularisant positif, on construit facilement des suites G_n, F_n et $\mathbf{q}_n$ de fonctions régulières sur $\bar{\Omega} \times [0,T]$ qui convergent vers Φ_δ, F et $\mathbf{q}$ dans $L^p(\Omega \times (0,T))$ si $p < \infty$ et telles que

$$\begin{aligned}
&\delta \leq G_n \leq M_\Phi \\
&|\mathbf{q}_n| \leq M_{\mathbf{q}} \\
&|F_n| \leq M_f.
\end{aligned}$$

Pour toute valeur de n, le théorème d'existence fournit une fonction ψ_n dans $\mathcal{C}$ vérifiant le problème dual associé à G_n, $\mathbf{q}_n$, F_n et χ. Ainsi, d'après (A.1)

$$\begin{aligned}
\int_0^T \int_\Omega u_d \chi &= \int_0^T \int_\Omega u_d(\psi_{nt} + G_n \Delta\psi_n + \mathbf{q}_n F_n.\nabla\psi_n) \\
&= \int_0^T \int_\Omega u_d(\psi_{nt} + \Phi\Delta\psi_n + \mathbf{q}F.\nabla\psi_n) + \int_0^T \int_\Omega u_d(G_n - \Phi)\Delta\psi_n \\
&\qquad + \int_0^T \int_\Omega u_d(\mathbf{q}_n F_n - \mathbf{q}F).\nabla\psi_n.
\end{aligned}$$

D'après (A.1),

$$\int_0^T \int_\Omega u_d(\psi_{nt} + \Phi\Delta\psi_n + \mathbf{q}F.\nabla\psi_n) = 0.$$

De plus, comme les majorations de G_n, $\mathbf{q}_n$, F_n et la minoration de G_n par δ sont indépendantes de n, les estimations sur $\Delta\psi_n$ et $\nabla\psi_n$ le sont également. Par conséquent, lorsque n tend vers l'infini,

$$\lim \int_0^T \int_\Omega u_d(\mathbf{q}_n F_n - \mathbf{q}F) \cdot \nabla\psi_n = 0$$

et

$$\limsup |\int_0^T \int_\Omega u_d(G_n - \Phi)\Delta\psi_n| \leq (\int_0^T \int_\Omega u_d^2(\Phi_\delta - \Phi)^2)^{\frac{1}{2}} \frac{C}{\delta^2},$$

car G_n converge vers Φ_δ et non vers Φ.

Mais $\Phi_\delta - \Phi \leq \delta\, 1\!\!1_{\{\Phi<\delta\}}$ puisque Φ et Φ_δ coincident sur $\{\Phi \geq \delta\}$. Donc si on note $A_\delta = \{u_d \neq 0\} \cap \{\Phi < \delta\}$, on obtient finalement

$$\int_0^T\int_\Omega u_d\chi \leq \Big(\int_{A_\delta} \frac{u_d^2}{\delta^2}\Big)^{\frac{1}{2}}.$$

Dans le cas général, le terme de droite ne tend pas forcément vers 0, car on n'a pas de controle de u_d sur A_δ. Cependant, si φ^{-1} est $\alpha - holder$,

$$\Phi(x,t) \leq \delta \Rightarrow |u_d(x,t)| \leq \delta^\alpha |u_d(x,t)|^\alpha.$$

Ainsi, pour $\alpha = \frac{1}{2}$, on a

$$|u_d| \leq \delta \text{ sur } A_\delta,$$

et donc

$$\int_0^T\int_\Omega u_d\chi \leq \sqrt{mes(A_\delta)}.$$

D'autre part, A_δ décroit vers $A_0 = \{u_d = 0\} \cap \{\Phi = 0\}$ si δ décroit vers 0 et A_0 est l'ensemble vide car φ est strictement croissante. Ainsi $mes(A_\delta) \to 0$ et

$$\int_0^T\int_\Omega u_d\chi = 0.$$

On termine la démonstration en prenant une suite χ_ε qui converge vers $sign(u_d)$ dans L^1, ce qui montre que $u_d = 0$ p.p. dans $\Omega \times (0,T)$.

Remark A.2.3 *Cette rpeuve permet donne une propriété de L^1-contraction pour le semi-groupe associé au problème. C'est à dire que si u et v sont deux solutions associées aux conditions initiales u_0 et v_0 alors*

$$\int_Q |u-v|dxdt \leq T\int_\Omega |u_0 - v_0|dx \tag{A.5}$$

A.3 Démonstration du théorème A.2.2

Pour démontrer ce théorème, on aura besoin de principes de maximum

Soit $Q_T = \Omega \times (0,T]$, et soit $\Gamma_T = \bar{Q}_T \backslash Q_T$. Si on pose $L_u = \Phi\Delta u + \mathbf{q}F.\nabla u$, le premier résultat très classique s'énonce ainsi :

Theorem A.3.1 (Principe du maximum intérieur) *Si $\theta \in C^{2,1}(\bar\Omega \times [0,T])$ vérifie $\theta_t \leq L_\theta$ dans Q_T alors*

$$\sup_{\bar{Q}_T} \theta = \max_{Q_T} \theta = \max_{\Gamma_T} \theta.$$

Démonstration.
Supposons dans un premier temps que $\theta_t < L_\theta$ dans Q_T et que θ atteint son maximum en un point P de Q_T. Dans ce cas, toutes les dérivées premières en espace de θ sont nulles et la hessienne de θ est négative au point P. En particulier, sa trace qui est $\Delta\theta$ est négative.
Ainsi,

$$L_\theta(P) \leq 0,$$

ce qui permet de dire que

$$\theta_t(P) < 0,$$

et donc que θ croit au voisinage de P le long d'une ligne verticale partant de P dans le sens des t décroissants. Cet argument contredit le fait que P soit un maximum.

Maintenant, si on a uniquement $\theta_t \leq L_\theta$, on introduit une fonction auxiliaire

$$\beta(x,t) = \exp(\lambda x_1).$$

On obtient alors facilement

$$L_\beta = \lambda\beta[\Phi\lambda + \mathbf{q}F.e_1],$$

donc si $\lambda > \dfrac{\sup \mathbf{q}F.e_1}{\delta}$, $L_\beta > 0$.

Ainsi, pour tout $\varepsilon > 0$, $L(\theta + \varepsilon\beta) > (\theta + \varepsilon\beta)_t$ et d'après ce qui vient d'etre fait,

$$\sup_{Q_T}(\theta + \varepsilon\beta) = \max_{\Gamma_T}(\theta + \varepsilon\beta).$$

En faisant tendre ε vers 0, on obtient la conclusion.

Etant donné que sur $\partial\Omega \times (0,T]$ on ne donne pas d'inégalités sur la valeur de ψ mais seulement sur sa dérivée normale, on aura besoin d'un second résultat.

Theorem A.3.2 (Principe du maximum sur le bord) *Si $\theta \in C^{2,1}(\bar{\Omega} \times [0,T])$ vérifie $\theta_t \leq L_\theta$ dans Q_T atteint son maximum de manière stricte en un point $P = (x_P, t_P)$ du " bord vertical" $\partial\Omega \times (0,T]$, au sens ou $\theta(S) < \theta(P)$ si $S \in Q_T$*
alors

$$\nabla u.\mathbf{n}(P) > 0$$

Démonstration
Dans un premier temps supposons que $t_P < T$. Comme dans la preuve du théorème précédent on construit une fonction auxiliaire, un peu plus compliquée cette fois ci, de manière à obtenir l'inégalité stricte sur la dérivée normale.

Soit Q un point de Q_T situé sur la normale sortante passant par P à Q_T et tel que l'adhérence de la boule B de centre Q passant par P privée du point P soit incluse dans Q_T. On note R le rayon de B. Comme $Q \neq P$, on peut construire une seconde boule B_0 de centre P telle que pour tout point de B_0, $\|x - x_Q\| \geq \alpha > 0$. On note $K = B \cap B_0$.

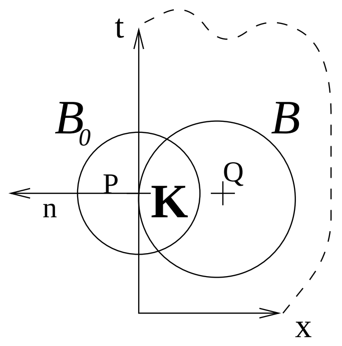


Fig 22. *Illustration dans le cas "1D en espace"*

On définit alors une fonction auxiliaire β sur K par

$$\beta(x,t) = e^{-\lambda(\|x-x_Q\|^2+(t-t_Q)^2)} - e^{-\lambda R^2}.$$

- *première propriété:* β est nulle sur ∂B.

- *seconde propriété:* $L_\beta - \beta_T > 0$ (si λ est assez grand).

En effet,

$$L_\beta - \beta_t = \lambda e^{-\lambda(\|x-x_Q\|^2+(t-t_Q)^2)}(4\Phi\lambda\|x-x_Q\|^2 - 2\mathbf{q}F.(x-x_Q,0) + 2(t-t_Q)).$$

Mais $\|x - x_Q\| \geq \alpha > 0$, donc si λ est suffisament grand, on a bien l'inégalité voulue.

La première partie de la démonstration du théorème précédent nous dit que si φ vérifie $\varphi_t - L_\varphi < 0$ sur un domaine D, elle ne peut atteindre son maximum que sur le bord de D.

Comme précédemment on construit γ_ε pour $\varepsilon > 0$

$$\gamma_\varepsilon = \theta + \varepsilon\beta.$$

Cette fonction vérifie bien $\gamma_{\varepsilon t} - L_{\gamma_\varepsilon} < 0$ sur K, par conséquent, elle atteint son maximum sur le bord de K.

Or $\gamma_\varepsilon = \theta$ sur ∂B. De plus, ∂K privé de ∂B est d'adhérence compacte dans Q_T et $\theta(S) < \theta(P)$ sur Q_T donc pour ε assez petit, on a également $\gamma_\varepsilon < \theta(P)$ sur ∂K privé de ∂B. De ces deux propriétés, on déduit que le maximum de γ_ε sur le bord de K est atteint en P.

Conclusion :

$$\theta(P) - \theta(S) \geq \varepsilon\beta(S), \quad \forall S \in K.$$

Si on divise les deux quantités par $d(S,P)$ et que l'on regarde la limite lorsque S tend vers P radialement (ie selon la droite (QP)), on obtient

$$\nabla\theta.\mathbf{n}(P) \geq 2\varepsilon\lambda R e^{(-\lambda R^2)} > 0,$$

ce qui termine la preuve dans le cas où $t_P < T$.
Lorsque $t_P = T$, la boule B n'est pas incluse dans Q_T. On doit regarder cette fois ci $K' = B \cap B_0 \cap Q_T$, la partie de la frontière de K' que l'on n'a pas prise en compte ($\{t = T\} \cap K$) ne fait pas partie de la

"frontière parabolique" de K' car pour chacun des points incriminés, on peut descendre verticalement sur une longueur strictement positive, en restant dans K'. Aussi la démonstration reste la même que pour les autres cas avec K' au lieu de K.

Démonstration du théorème A.2.2 :

En utilisant le principe du maximum intérieur, on est assuré qu'il est atteint sur les bords. D'autre part, la conclusion du principe du maximum sur le bord est fausse, ce qui implique que ses hypothèses ne sont pas vérifiées, et donc que le maximum est bien atteint sur le bord $\{t = 0\}$ où la fonction est par hypothèses négative. Ce qui termine la preuve du théorème A.2.2.

Bibliographie

[AA99] M Afif and B. Amaziane, *On convergence of finite volume schemes for one-dimensionnal two-phase flow in porous media*, Finite volumes for complex applications, Problems and Perspectives (F. Vilsmeier, R. Benkhaldoum and D. Hanël, eds.), Hermes, 1999, pp. 387–394.

[AdB85] H. W. Alt and E. di Benedetto, *Nonsteady flow of water and oil through inhomogeneous porous media*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **12** (1985), no. 3, 335–392. MR 87k:35043

[AL83] H. W. Alt and S. Luckhaus, *Quasilinear elliptic-parabolic differential equations*, Math. Z. **183** (1983), no. 3, 311–341. MR 85c:35059

[AS79] K. Aziz and A. Settari, *Petroleum reservoir simulation*, Applied Science Publishers, London, 1979.

[AWZ96] T. Arbogast, M. F. Wheeler, and N. Y. Zhang, *A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media*, SIAM J. Numer. Anal. **33** (1996), no. 4, 1669–1687. MR 97m:65172

[BCV95] S. Benharbit, A. Chalabi, and J.-P. Vila, *Numerical viscosity and convergence of finite volume methods for conservation laws with boundary conditions*, SIAM J. Numer. Anal. **32** (1995), no. 3, 775–796. MR 96c:65141

[Bea67] J. Bear, *Dynamic of flow in porous media*, Dover, 1967.

[Bea96] ———, *Modeling transport phenomena in porous media*, Environmental studies (Minneapolis, MN, 1992), Springer, New York, 1996, pp. 27–63. MR 1 417 048

[Ben72] P. Benilan, *Semi groupes d'évolution nonlinéaires*, Thèse de Doctorat de l'Université d'Orsay, Paris 11 (1972).

[BG95] P. Bénilan and R. Gariepy, *Strong solutions in L^1 of degenerate parabolic equations*, J. Differential Equations **119** (1995), no. 2, 473–502. MR 96c:35078

[BGN00] F. Bouchut, F. R. Guarguaglini, and R. Natalini, *Diffusive BGK approximations for nonlinear multidimensional parabolic equations*, Indiana Univ. Math. J. **49** (2000), no. 2, 723–749. MR 1 793 689

[BJ91] Yann Brenier and Jérôme Jaffré, *Upstream differencing for multiphase flow in reservoir simulation*, SIAM J. Numer. Anal. **28** (1991), no. 3, 685–696. MR 92i:76070

[BlRN79] C. Bardos, A. Y. le Roux, and J.-C. Nédélec, *First order quasilinear equations with boundary conditions*, Comm. Partial Differential Equations **4** (1979), no. 9, 1017–1034. MR 81b:35052

[Bré83] H. Brézis, *Analyse fonctionnelle*, Paris, 1983, Théorie et applications. [Theory and applications]. MR 85a:46001

[BT95] P. Bénilan and H. Touré, *Sur l'équation générale $u_t = a(\cdot, u, \varphi(\cdot, u)_x)_x + v$ dans L^1. I. Étude du problème stationnaire*, Evolution equations (Baton Rouge, LA, 1992), Dekker, New York, 1995, pp. 35–62. MR 95i:35135

[Car99] J. Carrillo, *Entropy solutions for nonlinear degenerate problems*, Arch. Ration. Mech. Anal. **147** (1999), no. 4, 269–361. MR 2000m:35132

[CCL95] B. Cockburn, F. Coquel, and P. G. LeFloch, *Convergence of the finite volume method for multidimensional conservation laws*, SIAM J. Numer. Anal. **32** (1995), no. 3, 687–705. MR 97f:65051

[CE99] Z. Chen and R. Ewing, *Mathematical analysis for reservoir models*, SIAM J. Math. Anal. **30** (1999), no. 2, 431–453 (electronic). MR 99m:76107

[CG99] B. Cockburn and G. Gripenberg, *Continuous dependence on the nonlinearities of solutions of degenerate parabolic equations*, J. Differential Equations **151** (1999), no. 2, 231–251. MR 2000a:35101

[CGH93] S. Champier, T. Gallouët, and R. Herbin, *Convergence of an upstream finite volume scheme for a nonlinear hyperbolic equation on a triangular mesh*, Numer. Math. **66** (1993), no. 2, 139–157. MR 95b:65117

[CH99] C. Chainais-Hillairet, *Finite volume schemes for a nonlinear hyperbolic equation. Convergence towards the entropy solution and error estimate*, M2AN Math. Model. Numer. Anal. **33** (1999), no. 1, 129–156. MR 2000a:65119

[Che97] Z. Chen, *Degenerate two phase flow incompressible flow 1 : existence, uniqueness and regularity of a weak solution*, SMU Math Report 97-10, Departement of Mathematics, Southern Methodist University (1997).

[Che99] ———, *Degenerate two phase flow incompressible flow 2 : regularity, stability and stabilization*, SMU Math Report 99-08, Departement of Mathematics, Southern Methodist University (1999).

[CJ86] G. Chavent and J. Jaffré, *Mathematical models and finite elements for reservoir simulation*, Elsevier, 1986.

[CP95] G. R. Cirmi and M. M. Porzio, *L^∞-solutions for some nonlinear degenerate elliptic and parabolic equations*, Ann. Mat. Pura Appl. (4) **169** (1995), 67–86. MR 96j:35088

[CVV99] Y. Coudière, J.P. Vila, and P. Villedieu, *Convergence rate of a finite volume scheme for a two-dimensional convection-diffusion problem*, M2AN Math. Model. Numer. Anal. **33** (1999), no. 3, 493–516. MR 1 713 235

[Dei85] Klaus Deimling, *Nonlinear functional analysis*, Springer-Verlag, Berlin, 1985. MR 86j:47001

[DiB93] E. DiBenedetto, *Degenerate parabolic equations*, Springer-Verlag, New York, 1993. MR 94h:35130

[DiP85] R. J. DiPerna, *Measure-valued solutions to conservation laws*, Arch. Rational Mech. Anal. **88** (1985), no. 3, 223–270. MR 86g:35121

[Dro] J. Droniou, *Ouverts polygonaux à frontiere lipschitzienne.*

[EGGH98] R. Eymard, T. Gallouët, M. Ghilani, and R. Herbin, *Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by finite volume schemes*, IMA J. Numer. Anal. **18** (1998), no. 4, 563–594. MR 2000b:65180

[EGH95a] R. Eymard, T. Gallouët, and R. Herbin, *Convergence of a finite volume scheme for a nonlinear hyperbolic equation*, Proceedings of the Third International Colloquium on Numerical Analysis (Plovdiv, 1994) (Utrecht), VSP, 1995, pp. 61–70. MR 1 455 950

[EGH95b] ______, *Existence and uniqueness of the entropy solution to a nonlinear hyperbolic equation*, Chinese Ann. Math. Ser. B **16** (1995), no. 1, 1–14, A Chinese summary appears in Chinese Ann. Math. Ser. A **16** (1995), no. 1, 119. MR 96m:35202

[EGH99] ______, *Convergence of finite volume schemes for semilinear convection diffusion equations*, Numer. Math. **82** (1999), no. 1, 91–116. MR 2000b:65213

[EGH00a] ______, *Error estimate of approximate solutions of a nonlinear convection-diffusion problem*, Prépublication LATP (2000).

[EGH00b] R. Eymard, T. Gallouët, and R Herbin, *The finite volume method*, Handbook for Numerical Analysis, Ph. Ciarlet J.L. Lions eds, North Holland, Paris, 2000, pp. 715–1022.

[EGHM00] R. Eymard, T. Gallouet, R. Herbin, and A. Michel, *Convergence of a finite volume scheme for parabolic degenerate equations*, accepté pour publication dans Numer. Math. (2000).

[EGHNS98] R. Eymard, T. Gallouët, D. Hilhorst, and Y. Naït Slimane, *Finite volumes and nonlinear diffusion equations*, RAIRO Modél. Math. Anal. Numér. **32** (1998), no. 6, 747–761. MR 99j:65166

[EHM01] R. Eymard, R. Herbin, and A. Michel, *Mathematical study of a petroleum engeneering scheme*, En préparation (2001).

[EK99] S. Evje and K. H. Karlsen, *Degenerate convection-diffusion equations and implicit monotone difference schemes*, Hyperbolic problems: theory, numerics, applications, Vol. I (Zürich, 1998), Birkhäuser, Basel, 1999, pp. 285–294. MR 2000k:65137

[EK00a] ______, *Discrete approximations of BV solutions to doubly nonlinear degenerate parabolic equations*, Numer. Math. **86** (2000), no. 3, 377–417. MR 1 785 415

[EK00b] ______, *Monotone difference approximations of BV solutions to degenerate convection-diffusion equations*, SIAM J. Numer. Anal. **37** (2000), no. 6, 1838–1860 (electronic). MR 1 766 850

[EW84] R. E. Ewing and M. F. Wheeler, *Galerkin methods for miscible displacement problems with point sources and sinks—unit mobility ratio case*, Mathematical methods in energy research (Laramie, Wyo., 1982/1983), SIAM, Philadelphia, PA, 1984, pp. 40–58. MR 86k:76060

[Fai92] I. Faille, *A control volume method to solve an elliptic equation on a two-dimensionnal irregular meshing*, Comp. Meth. Appl. Mech. Engrg. **100** (1992), 275–290.

[FG00] P. Fabrie and T. Gallouët, *Modeling wells in porous media flow*, Math. Models Methods Appl. Sci. **10** (2000), no. 5, 673–709. MR 1 763 198

[GHM99] S. Gerbi, R. Herbin, and E. Marchand, *Existence of a solution to a coupled elliptic system with a Signorini condition*, Adv. Differential Equations **4** (1999), no. 2, 225–250. MR 2000a:35070

[GMT96] G. Gagneux and M. Madaune-Tort, *Analyse mathématique de modèles non linéaires de l'ingénierie pétrolière*, Springer-Verlag, Berlin, 1996, With a preface by Charles-Michel Marle. MR 1 616 513

[GT77] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin, 1977, Grundlehren der Mathematischen Wissenschaften, Vol. 224. MR 57 13109

[Her96] R. Herbin, *Finite volume methods for diffusion convection equations on general meshes*, Finite volumes for complex applications, Problems and Perspectives (F. Benkhaldoun and R. Vilsmeier, eds.), Hermes, 1996.

[KL84] D. Kroener and S. Luckhaus, *Flow of oil and water in a porous medium*, J. Differential Equations **55** (1984), no. 2, 276–288. MR 85k:35211

[KO99] D. Kröner and M. Ohlberger, *A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multi-dimensions*, Math. Comp. (1999). MR 1 659 843

[Kru70] S. N. Kruzkov, *First order quasilinear equations with several independent variables.*, Mat. Sb. (N.S.) **81 (123)** (1970), 228–255. MR 42 2159

[LBS93] A. Lagha-Benadallah and F. Smadhi, *Existence de solutions faibles pour un problème aux limites associé à une équation parabolique dégénérée*, Rev. Maghrébine Math. **2** (1993), no. 2, 201–221. MR 97d:35112

[LSU67] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Uralceva, *Linear and quasilinear equations of parabolic type*, American Mathematical Society, Providence, R.I., 1967, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. MR 39 3159b

[Mic99] A. Michel, *Convergence of a finite volume scheme for a nonlinear convection-diffusion problem*, Finite volume methods for complex applications II, Problems and Prerspectives (R. Vilsmeier, F. Benkhaldoum, and D. Hanel, eds.), Hermes, 1999, pp. 173–180.

[Mic01] ______, *A finite volume scheme for the simulation of two-phase incompressible flow in porous media*, Submitted. (2001).

[MNRR96] J. Malek, J. Necas, M. Rokyta, and M. Ruzicka, *Weak and measure-valued solutions to evolutionary PDEs*, Chapman & Hall, London, 1996. MR 97g:35002

[MPT00] C. Mascia, A. Porrata, and A. Terracina, *Entropy solutions of boundary value problems for degenerate parabolic-hyperbolic equations*, Prépublication (2000).

[MV01] A. Michel and J. Vovelle, *A finite volume method for parabolic degenerate problems with general dirichlet boundary conditions*, En préparation. (2001).

[NS97] Y. Nait Slimane, *Méthodes de volumes finis pour des problèmes de diffusion-convection non linéaires*, Thèse de Doctorat de l'Université Paris 13 (1997).

[Ohl97] M. Ohlberger, *Convergence of a mixed finite elements–finite volume method for the two phase flow in porous media*, East-West J. Numer. Math. **5** (1997), no. 3, 183–210. MR 98m:76105

[Ott96a] F. Otto, *Initial-boundary value problem for a scalar conservation law*, C. R. Acad. Sci. Paris Sér. I Math. **322** (1996), no. 8, 729–734. MR 97c:35132

[Ott96b] Felix Otto, *L^1-contraction and uniqueness for quasilinear elliptic-parabolic equations*, J. Differential Equations **131** (1996), no. 1, 20–38. MR 97i:35125

[PDDGU97] A. Plouvier-Debaigt, B. Donné, G. Gagneux, and P. Urruty, *Solutions renormalisées pour des modèles des milieux poreux*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 10, 1091–1095. MR 99b:35109

[PDG96] A. Plouvier-Debaigt and G. Gagneux, *Unicité de la solution faible d'équations des milieux poreux, via un concept de solution renormalisée*, C. R. Acad. Sci. Paris Sér. I Math. **322** (1996), no. 11, 1049–1052. MR 97h:35075

[Por96] M. M. Porzio, *L^∞_{loc} estimates for a class of doubly nonlinear parabolic equations with sources*, Rend. Mat. Appl. (7) **16** (1996), no. 3, 433–456. MR 97i:35080

[Por97] ———, *L^∞-regularity for degenerate and singular anisotropic parabolic equations*, Boll. Un. Mat. Ital. A (7) **11** (1997), no. 3, 697–707. MR 98g:35115

[Por99] ———, *Existence of solutions for some "noncoercive" parabolic equations*, Discrete Contin. Dynam. Systems **5** (1999), no. 3, 553–568. MR 2000c:35107

[PV93] M. M. Porzio and V. Vespri, *Hölder estimates for local solutions of some doubly nonlinear degenerate parabolic equations*, J. Differential Equations **103** (1993), no. 1, 146–178. MR 94d:35015

[RG99] E. Rouvre and G. Gagneux, *Solution forte entropique de lois scalaires hyperboliques-paraboliques dégénérées*, C. R. Acad. Sci. Paris Sér. I Math. **329** (1999), no. 7, 599–602. MR 1 717 116

[Ser96] D. Serre, *Systèmes de lois de conservation*, Diderot, 1996.

[Sze89] A. Szepessy, *An existence result for scalar conservation laws using measure valued solutions*, Communication in Partial Differential Equations (1989).

[Vig96] M. H. Vignal, *Convergence of a finite volume scheme for an elliptic-hyperbolic system*, RAIRO Modél. Math. Anal. Numér. **30** (1996), no. 7, 841–872. MR 98a:65143

[Vil94] J.-P. Vila, *Convergence and error estimates in finite volume schemes for general multidimensional scalar conservation laws. I. Explicit monotone schemes*, RAIRO Modél. Math. Anal. Numér. **28** (1994), no. 3, 267–295. MR 96d:65150

[Vov00] J. Vovelle, *Convergence of finite volume monotone scheme for scalar conservation law on bounded domain*, accepté pour publication dans Numer. Math. (2000).

Résumé. Ce mémoire est centré autour de l'analyse numérique de schémas volumes finis pour un modèle simplifié d'écoulement de deux fluides incompressibles en milieu poreux. Ces phénomènes sont souvent qualifiés de phénomènes de convection diffusion à convection dominante ("convection dominated problems" en anglais).

La première partie du mémoire est consacrée à l'approximation numérique d'équations paraboliques hyperboliques faiblement ou fortement dégénérées. Les trois premiers chapitres sont consacrés à l'étude de la convergence de schémas volumes finis. Le dernier chapitre est consacré à l'analyse des résultats numériques obtenus.

La seconde partie est consacrée à l'analyse numérique d'un modèle simplifié d'écoulement diphasique en milieu poreux par deux schémas différents. Le premier schéma dit "des mathématiciens" est basé sur la réécriture du système étudié sous la forme d'une équation parabolique hyperbolique sur la saturation et d'une équation elliptique sur la pression, ces deux équations étant couplées par le coefficient de diffusion. Le second schéma dit schéma "des pétroliers" est une méthode numérique utilisée en pratique dans l'industrie pétrolière. Les deux schémas sont analysés séparément et ils sont ensuite comparés numériquement.

Mots clés : milieux poreux, écoulements diphasiques incompressibles, schémas numériques, volumes finis, équations paraboliques dégénérées, équations hyperboliques, formulation entropique, estimations a priori, convergence, unicité, mesures d'Young, Kruzkov.

Discipline : mathématiques

Convergence of finite volumes schemes for nonlinear convection diffusion problems

Abstract. The main subject of this report is the mathematical study of finite volume schemes for problems of two phase flow in porous media. These problems are often called "convection dominated problems".

The first part concerns the numerical approximation of degenerate hyperbolic parabolic equations by a finite volume method. In Chapter 1, Chapter 2 and Chapter 3, We study the convergence of the scheme and in Chapter 4, we present numerical results.

The second part of this report concerns the study of two numerical methods for a simplified model of two phase flow in porous media. For the first scheme also called "schéma des mathématiciens", we transform the coupled system in an equivalent system of two equations, a parabolic equation on the saturation and an elliptic equation on the pressure coupled by the convective term. The second scheme is a phase by phase upwinding finite volume scheme which is used in petroleum industry. We study the convergence of the schemes and we make a comparison between the two methods by using numerical experiments.

Key words : porous media, two phase incompressible flow, numerical schemes, finite volumes, degenerate parabolic equations, hyperbolic equations, entropic formulation, a priori estimates, convergence, uniqueness, Young measure, Kruzkov.

Laboratoire d'Analyse, Topologie et Probabilités, C.M.I. Marseille, 39 rue F.Joliot Curie, 13453 Marseille Cedex 13.